ADHOC NETWORKS

Sub Code	**: 10IS841/10CS841**	**IA Marks**	**: 25**
Hrs/Week	**: 04**	**Exam Hours**	**: 03**
Total Hrs	**: 52**	**Exam Marks**	**: 100**

PART – A

UNIT 1 **6 Hours**
Introduction: Ad hoc Networks: Introduction, Issues in Ad hoc wireless networks, Ad hoc wireless internet.

UNIT 2 **7 Hours**
MAC – 1: MAC Protocols for Ad hoc wireless Networks: Introduction, Issues in designing a MAC protocol for Ad hoc wireless Networks, Design goals of a MAC protocol for Ad hoc wireless Networks, Classification of MAC protocols, Contention based protocols with reservation mechanisms.

UNIT 3 **6 Hours**
MAC – 2: Contention-based MAC protocols with scheduling mechanism, MAC protocols that use directional antennas, Other MAC protocols.

UNIT 4 **7 Hours**
Routing – 1: Routing protocols for Ad hoc wireless Networks: Introduction, Issues in designing a routing protocol for Ad hoc wireless Networks, Classification of routing protocols, Table drive routing protocol, On-demand routing protocol.

PART- B

UNIT 5 **6 Hours**
Routing – 2: Hybrid routing protocol, Routing protocols with effective flooding mechanisms, Hierarchical routing protocols, Power aware routing protocols

UNIT 6 **7 Hours**
Transport Layer: Transport layer protocols for Ad hoc wireless Networks: Introduction, Issues in designing a transport layer protocol for Ad hoc wireless Networks, Design goals of a transport layer protocol for Ad hoc wireless Networks, Classification of transport layer solutions, TCP over Ad hoc wireless Networks, Other transport layer protocols for Ad hoc wireless Networks.

UNIT 7 **6 Hours**
Security: Security: Security in wireless Ad hoc wireless Networks, Network security requirements, Issues & challenges in security provisioning, Network security attacks, Key management, Secure routing in Ad hoc wireless Networks.

UNIT 8 **7 Hours**
QoS: Quality of service in Ad hoc wireless Networks: Introduction, Issues and challenges in providing QoS in Ad hoc wireless Networks, Classification of QoS solutions, MAC layer solutions, network layer solutions.

Text Books:

1. C. Siva Ram Murthy & B. S. Manoj: Ad hoc Wireless Networks, 2nd Edition, Pearson Education, 2005

TABLE OF CONTENTS

UNIT 1: INTRODUCTION

CELLULAR AND ADHOC WIRELESS NETWORKS

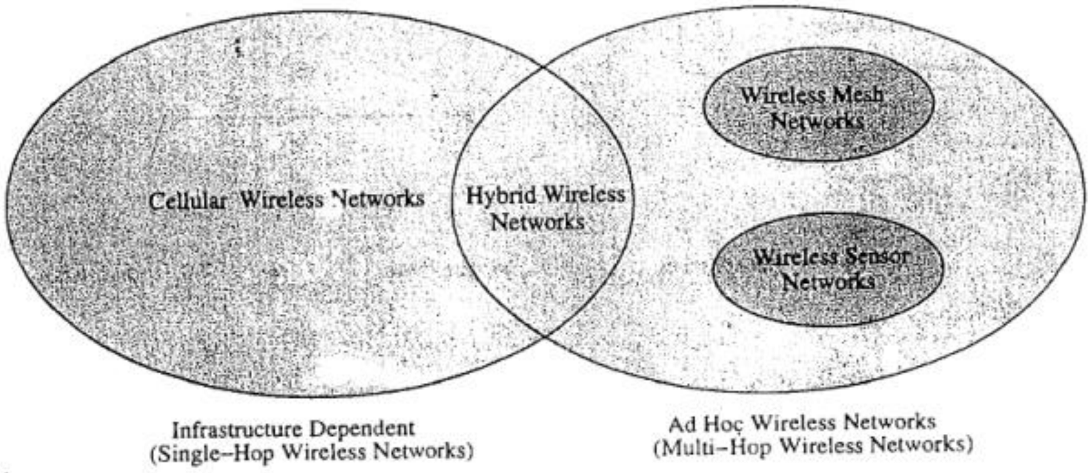

Cellular Wireless Networks
Hybrid Wireless Networks
Wireless Mesh Networks
Wireless Sensor Networks
Infrastructure Dependent (Single-Hop Wireless Networks)
Ad Hoc Wireless Networks (Multi-Hop Wireless Networks)

Figure 5.1: Cellular and adhoc wireless networks

- The current cellular wireless networks are classified as the infrastructure-dependent network.
- The path-setup for a call between two nodes, say, node C to E, is completed through base-station(Fig 5.2)

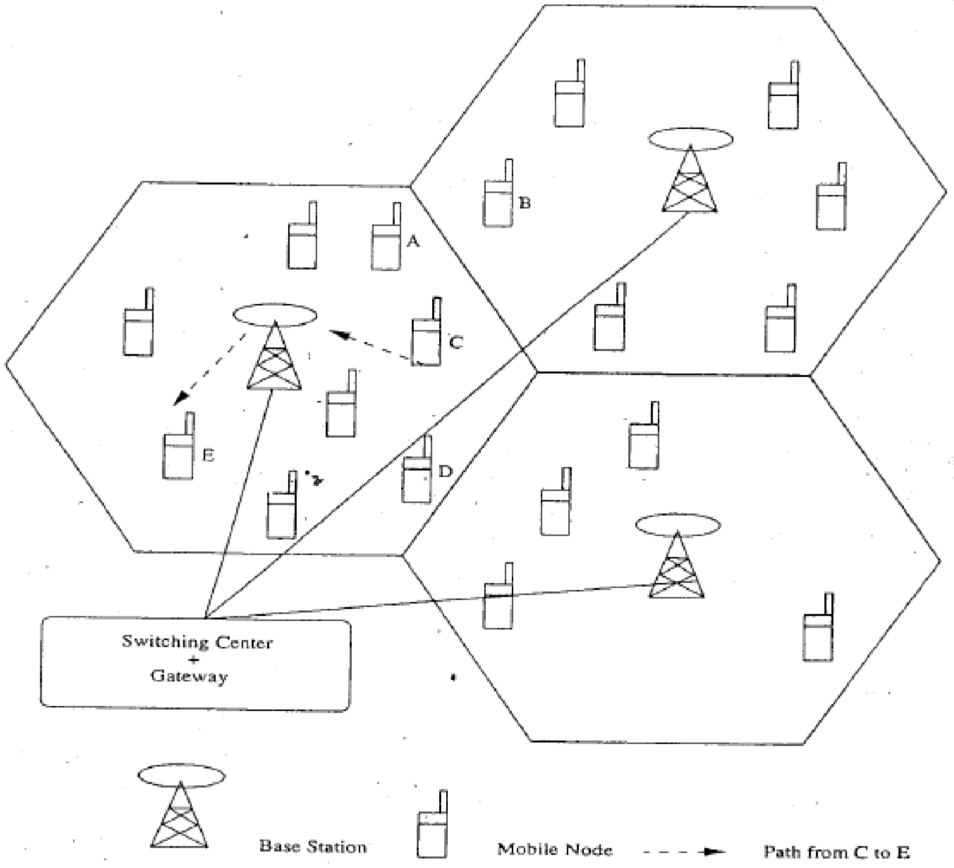


Figure 5.2: A cellular network

ADHOC NETWORKS

- Adhoc wireless networks are defined as a category of wireless network.
- They utilize multi-hop radio replaying.
- They are capable of operating without the support of any fixed infrastructure.
- Absence of any central co-ordinator (or base station) makes the routing complex.
- Adhoc network topology for the cellular network is illustrated below (Fig 5.3).
- The path-setup for a call between 2 nodes, say, node C to E, is completed through the intermediate mobile node F.

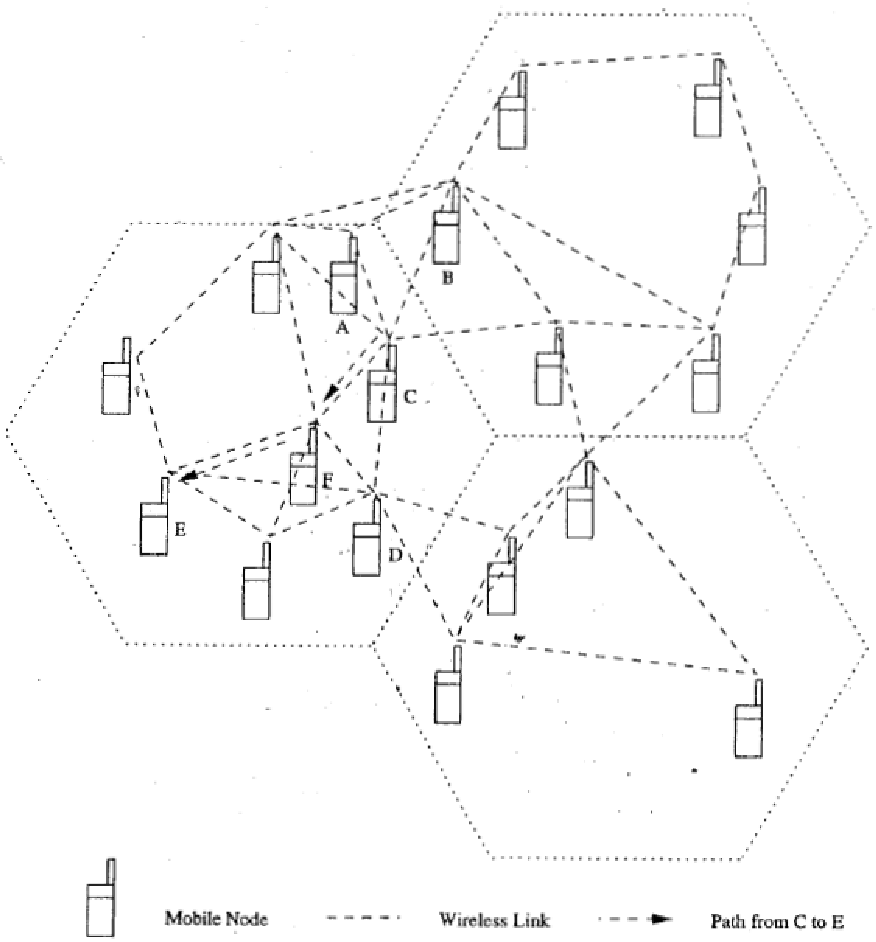
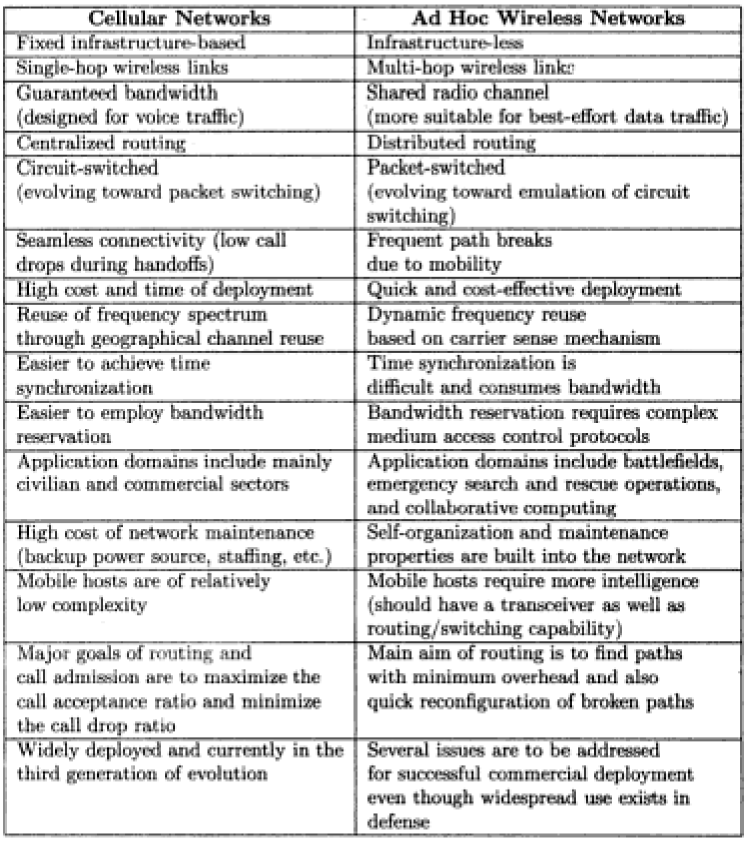


Figure 5.3: An adhoc wireless network

Difference between Cellular Networks and Adhoc Wireless Networks



Cellular Networks	Ad Hoc Wireless Networks
Fixed infrastructure-based	Infrastructure-less
Single-hop wireless links	Multi-hop wireless links
Guaranteed bandwidth (designed for voice traffic)	Shared radio channel (more suitable for best-effort data traffic)
Centralized routing	Distributed routing
Circuit-switched (evolving toward packet switching)	Packet-switched (evolving toward emulation of circuit switching)
Seamless connectivity (low call drops during handoffs)	Frequent path breaks due to mobility
High cost and time of deployment	Quick and cost-effective deployment
Reuse of frequency spectrum through geographical channel reuse	Dynamic frequency reuse based on carrier sense mechanism
Easier to achieve time synchronization	Time synchronization is difficult and consumes bandwidth
Easier to employ bandwidth reservation	Bandwidth reservation requires complex medium access control protocols
Application domains include mainly civilian and commercial sectors	Application domains include battlefields, emergency search and rescue operations, and collaborative computing
High cost of network maintenance (backup power source, staffing, etc.)	Self-organization and maintenance properties are built into the network
Mobile hosts are of relatively low complexity	Mobile hosts require more intelligence (should have a transceiver as well as routing/switching capability)
Major goals of routing and call admission are to maximize the call acceptance ratio and minimize the call drop ratio	Main aim of routing is to find paths with minimum overhead and also quick reconfiguration of broken paths
Widely deployed and currently in the third generation of evolution	Several issues are to be addressed for successful commercial deployment even though widespread use exists in defense

APPLICATIONS OF ADHOC WIRELESS NETWORKS

Military Application

- Adhoc networks can be used to establish communication among a group of soldiers for tactical operations.
- Setting up of a fixed infrastructure for communication among group of soldiers in enemy territories (or in inhospitable terrains) may not be possible. In such a case, adhoc networks can be used to provide required communication quickly.
- The military application enforces following requirements

* Reliability	* Efficiency
* Secure communication	* Support for multicast routing

Collaborative & Distributed Computing

- Adhoc network can be used to establish temporary communication infrastructure among a group of people in a conference.
- In distributed file sharing application, reliability is of high importance which would be provided by adhoc network.
- Other applications such as streaming of multimedia objects require support for soft real-time communication.
- Devices used can be

→ laptops with add-on wireless interface cards
→ enhanced PDA (Personal Digital Assistant) or
→ mobile devices with high processing power

Emergency Operations

• Adhoc networks can be used in emergency operations such as

* search and rescue
* crowd control &
* commando operations

• The major factors that favor adhoc networks for such tasks are

→ self-configuration of system with minimal overhead
→ independent of fixed infrastructure
→ freedom and flexibility of mobility &
→ unavailability of conventional communication infrastructure

• In environments where the infrastructure-based communication facilities are destroyed due to natural calamities (or due to a war), immediate deployment of adhoc networks would be a good solution for co-ordinating rescue activities.

Wireless Mesh Network (WMN)

• WMN can be formed to provide an alternate communication-infrastructure for mobile or fixed nodes,

* without the spectrum reuse constraint &
* without the requirement of network planning of cellular network

• It provides many alternate paths for data transfer between source & destination, which results in quick reconfiguration of path when the existing path fails due to node-failure.

• Major advantages are

* low cost of deployment
* high scalability
* support for a high data-rate
* easy extendability
* high availability &
* low cost/bit

• Since the infrastructure-built is in the form of small radio relaying-devices, the investment required is much less when compared to cellular-network counterpart.

• The possible deployment scenarios include

* residential zones
* highways
* business zones
* important civilian regions &
* university campuses

• It should be capable of self-organization and maintenance.

• It operates at license-free ISM band around 2.4 GHz & 5 GHz.

Wireless Sensor Networks (WSN)

- These are used to provide a wireless communication infrastructure among the sensors deployed in a specific application domain.
- Sensor-nodes are tiny devices that have capability of
  → sensing physical parameters
  → processing the data gathered &
  → communicating to the monitoring system
- ***The issues that make sensor network a distinct category of adhoc network are the following***

1. **Node Mobility**
   - Mobility of nodes is not a mandatory requirement in sensor-networks.
   - For example,
     → the nodes used for periodic monitoring of soil properties are not required to be mobile
     → the nodes fitted on the bodies of patients are designed to support partial mobility
   - In general, sensor-networks need not in all cases be designed to support mobility of nodes.
2. **Network Size**
   - The number of nodes in sensor network can be much larger than that in a typical adhoc network
3. **Density of Deployment**
   - The density of nodes varies with the domain of application.
   - For example, military applications require high availability of network, which makes redundancy a high priority.
4. **Power Constraints**
   - The power constraints in sensor networks are much more severe than those in adhoc networks. This is mainly because the nodes are expected to operate in harsh environmental conditions, with minimum human supervision & maintenance.
   - In certain case, the recharging of the energy source is impossible.
   - Running such a network demands very efficient protocol at network, data link, and physical layer.
   - The power sources can be classified into following 3 categories
     → *Replenishable Power Source:* The power source can be replaced when the existing source is fully drained.
     → *Non-replenishable Power Source:* The power source cannot be replenished once the network has been deployed. The replacement of node is the only solution.
     → *Regenerative Power Source:* Here, power source have the capability of regenerating power from the physical parameter under measurement.
5. **Data/Information Fusion**
   - Data fusion refers to the aggregation of multiple packets into one before relaying it.
   - Data fusion is used
     → to reduce bandwidth consumed by redundant headers of packets &
     → to reduce delay involved in transmitting multiple packets
   - Information fusion is used
     → to process sensed data at intermediate nodes &
     → to relay the outcome to the monitoring system
6. **Traffic Distribution**
   - The communication traffic pattern varies with the domain of application.
   - For example,
     → Environmental sensing application generates short periodic packets indicating the status of environmental parameter. This kind of traffic requires low bandwidth
     → Military applications generally carry user traffic such as digitized & packetized voice stream This kind of traffic requires high bandwidth.

Hybrid Wireless Networks

- One of the major application areas of adhoc network is in the hybrid wireless architecture such as Multi-hop Cellular Network[MCN] & Integrated Cellular Adhoc Relay[iCAR].
- The primary concept behind cellular networks is geographical channel-reuse.
- Several techniques like cell sectoring, cell resizing and multi tier cells increase the capacity of cellular networks.
- MCNs combine
  → reliability & support of fixed base-station of cellular network with
  → flexibility & multi-hop relaying adhoc networks
- Major advantages are as follows
  → Higher capacity than cellular networks due to the better channel reuse
  → Increased flexibility & reliability in routing
  → Better coverage & connectivity in holes of a cell can be provided by means of multiple hops through intermediate nodes in a cell

ISSUES IN AD HOC WIRELESS NETWORKS

The major issues that affect the design, deployment, & performance of an ad hoc wireless network system are:

- □ Medium Access Scheme
- □ Routing
- □ Multicasting
- □ Scalability
- □ Pricing Scheme
- □ QoSProvisioning
- □ Transport Layer Protocol
- □ Energy Management
- □ Self-Organization
- □ Addressing &Service discovery
- □ Deployment considerations

Medium Access Scheme

The primary responsibility of a Medium Access Control (MAC) protocol is:

the distributed arbitration for the shared channel for transmission of packets.

The major issues to be considered in designing a MAC protocol are as follows

1. Distributed Operation

• The adhoc networks need to operate in environments where no centralized co-ordination is possible.

• The MAC protocol design should be fully distributed involving minimum control-overhead.

2. Synchronization

• The MAC protocol design should take into account the requirement of time-synchronization.

• Synchronization is mandatory for TDMA-based systems for management of transmission & reception slots.

3. Hidden Terminals

• Hidden terminals are nodes that are hidden (or not reachable) from the sender of a data transmission session, but are reachable to the receiver of the session.

4. Exposed Terminals

• Exposed terminals, the nodes that are in the transmission-range of the sender of an on-going session, are prevented from making a transmission.

5. Throughput

• The MAC protocol should attempt to maximize the throughput of the system, which can be done by

→ Minimizing the occurrence of collisions.
→ Maximizing channel utilization and
→ Minimizing control overhead.

6. Access Delay

• It refers to the average delay that any packet experiences to get transmitted.

• The MAC protocol should attempt to minimize the delay.

7. Fairness

• It refers to the ability of the MAC protocol to provide an equal share (or weighted share) of the bandwidth to all competing nodes.

• Fairness can be either node-based or flow-based.

8. Resource Reservation

• The provisioning of QoS requires reservation of resources such as bandwidth, buffer space, and processing power.

9. Ability to measure Resource Availability

• The MAC protocol should be able to provide an estimation of resource availability at every node in order

→ to handle the resources efficiently and
→ to perform call admission control

• This can also be used for making congestion control decisions.

10. Capability for Power Control

• The transmission power control

→ reduces the energy consumption at the nodes
→ causes a decrease in interference at neighboring nodes &
→ increases frequency reuse

10. Adaptive Rate Control

• This refers to the variation in the data-rate achieved over a channel.

• The MAC protocol should make use of a high data-rate when the sender and receiver are nearby &
adaptively reduce the data-rate as they move away from each other.

11. Use of Directional Antennas

• This has many advantages that include

* Increased spectrum reuse
* Reduction in interference &
* Reduced power consumption

Routing

The responsibilities of a routing protocol include

→ exchanging the route information;
→ finding a feasible path to a destination.

The major challenges that a routing protocol faces are as follows:

1. Mobility

• The mobility of nodes results in

* frequent path breaks
* packet collisions
* transient loops
* stale routing information
* difficulty in resource reservation

2. Bandwidth Constraint

• Since the channel is shared by all nodes in the broadcast region, the bandwidth available per wireless link depends on the number of nodes & traffic they handle.

3. Error-prone & Shared Channel

• The Bit Error Rate (BER) in a wireless channel is very high [10^{-5} to 10^{-3}] compared to that in its wired counterparts [10^{-12} to 10^{-9}].

• Consideration of the state of the wireless link, signal-to-noise ratio, and path loss for routing in adhoc networks can improve the efficiency of the routing protocol.

4. Location Dependent Contention

• The load on the wireless channel varies with the number of nodes present in a given geographical region.

• When the number of nodes increases, the contention for the channel increases, which results in a high number of collisions & a subsequent wastage of bandwidth.

5. Other Resource Constraints

• The constraints on following resources limit the capability of a routing protocol

* computing power
* battery power &
* buffer storage

The major requirements of a routing protocol in adhoc networks are the following.

1. Minimum Route Acquisition Delay

• The route acquisition delay for a node that does not have a route to a particular destination node should be as minimal as possible.

• The delay may vary with the network-size and the network-load.

2. Quick Route Re-configuration

• The unpredictable changes in the network topology require that the routing protocol be able to quickly perform route re-configuration in order to handle path breaks and subsequent packet losses.

3. Loop-Free Routing

• This is a fundamental requirement to avoid unnecessary wastage of bandwidth.

• In adhoc networks, due to the random movement of nodes, transient loops may form in the route thus established.

• A routing protocol should detect such transient routing loops & take corrective actions.

4. Distributed Routing Approach

• An adhoc network is a fully distributed wireless network & the use of centralized routing approaches in such a network may consume a large amount of bandwidth.

5. Minimum Control Overhead

• The control packets exchanged for finding a new route, and maintaining existing routes should be kept as minimal as possible.

6. Scalability

• Scalability is the ability of the routing protocol to scale well in a network with a large number of nodes.

• This requires minimization of control overhead & adaptation of the routing protocol to the network size.

7. Provisioning of QoS

• The routing protocol should be able to provide a certain level of QoS as demanded by the nodes(or the category of calls).

• The QoS parameters can be bandwidth, delay, jitter, packet delivery ratio, & throughput.

8. Support for Time-sensitive Traffic

• Tactical communications & similar applications require support for time-sensitive traffic.

• The routing protocol should be able to support both hard real-time & soft real-time traffic.

9. Security & Privacy

• The routing protocol must be resilient to threats and vulnerabilities.

• It must have inbuilt capability to avoid attacks such as resource consumption, DOS, impersonation.

Multicasting

- It plays important role in emergency search-&-rescue operations & in military communication.
- Use of single-link connectivity among the nodes in a multicast group results in a tree-shaped multicast routing topology.
- Such a tree-shaped topology provides high multicast efficiency, with low packet delivery-ratio due to the frequency tree breaks.

The major issues in designing multicast routing protocols are as follows

1. Robustness

• The multicast routing protocol must be able to recover-&-reconfigure quickly from potential link breaks thus making it suitable for use in high dynamic environments.

2. Efficiency

•The protocol should make a minimum number of transmissions to deliver a data-packet to all the group-members.

3. Control Overhead

• The scarce bandwidth availability demands minimal control-overhead for the multicast session.

4. QoS

• QoS support is essential in multicast-routing because, in most cases, the data transferred in a multicast session is time-sensitive.

5. Efficient Group Management

• Group management refers to the process of

→ accepting multicast session members &
→ maintaining the connectivity among them until the session expires.

6. Scalability

• The protocol should be able to scale for a network with a large number of nodes.

7. Security

• Authentication of session-members and prevention of non-members from gaining unauthorized information play a major role in military communications.

Transport Layer Protocol

- The main objectives of the transport layer protocols include:
  * Setting-up & maintaining end-to-end connections
  * Reliable end-to-end delivery of packets
  * Flow control &
  * Congestion control
- Examples of some transport layers protocols are

a. UDP (User Datagram Protocol)

→ It is an unreliable connectionless transport layer protocol
→ It neither performs flow control & congestion control
→ It do not take into account the current network status such as congestion at the intermediate links, the rate of collision, or other similar factors affecting the network throughput

b. TCP (Transmission Control Protocol)

→ It is a reliable connection-oriented transport layer protocol
→ It performs flow control & congestion control
→ Here, performance degradation arises due to frequent path breaks, presence of stale routing information, high channel error rate, and frequent network partitions

Pricing Scheme

• Assume that an optimal route from node A to node B passes through node C, & node C is not powered on.
• Then node A will have to set up a costlier & non-optimal route to B.
• The non-optimal path consumes more resources & affects the throughput of the system.
• As the intermediate nodes in a path that relay the data packets expend their resources such as battery-charge & computing-power, they should be properly compensated.
• Hence, pricing schemes that incorporate service compensation (or service reimbursement) are required.

Self-Organization

• One very important property that an adhoc network should exhibit is organizing & maintaining the network by itself.
• The major activities are

→ Neighbor discovery
→ Topology organization &
→ Topology reorganization (updating topology information)

QoS Provisioning

• QoS is the performance level of services offered by a service provider (or a network) to the user.

• QoS provisioning often requires

→ Negotiation between host & the network
→ Resource reservation schemes
→ Priority scheduling &
→ Call admission control

QoS Parameters

Applications	*Corresponding QoS parameter*
1. Multimedia application ->	1.Bandwidth & Delay
2. Military application ->	2.Security & Reliability
3. Defense application ->	3. Finding trustworthy intermediate hosts & routing.
4. Emergency search and rescue operations ->	4.Availability.
5. Communication among the nodes in a sensor network ->	5.Minimum energy consumption

QoS-Aware Routing

i. Finding the path is the first step toward a QoS-aware routing protocol.

ii. The parameters that can be considered for routing decisions are

* Network throughput
* Reliability
* Delay jitter
* Bit error rate
* Packet delivery ratio
* Delay
* Packet loss rate
* Path loss

QoS Framework

- A framework for QoS is a complete system that attempts to provide the promised services to each user or application.
- The key component of QoS framework is a QoS service model which defines the way user requirements are served.

Addressing & Service Discovery

- Addressing & service discovery assume significance in adhoc network due to the absence of any centralized coordinator.
- An address that is globally unique in the connected part of the adhoc network is required for a node in order to participate in communication.
- Auto-configuration of addresses is required to allocate non-duplicate addresses to the nodes.

Scalability

- Scalability is the ability of the routing protocol to scale well in a network with a large number of nodes.
- It requires minimization of control overhead & adaptation of the routing protocol to the network size.

Security

1) Security is an important issue in adhoc network as the information can be hacked.

2) Attacks against network are of 2 types:

I. Passive attack → Made by malicious node to obtain information transacted in the network without disrupting the operation

II. Active attack → They disrupt the operation of network

Further active attacks are of 2 types:

- External attack: The active attacks that are executed by nodes outside the network
- Internal attack: The active attacks that are performed by nodes belonging to same network

3) The major security threats that exist in adhoc networks are as follows

- **DoS (Denial of Service)** – The attack affected by making the network resource unavailable for service to other nodes, either
  → by consuming the bandwidth or
  → by overloading the system
- **Resource Consumption** – The scarce availability of resources in adhoc network makes it an easy target for internal attacks, particularly aiming at consuming resources available in the network. The major types of resource consumption attacks are,
  I. **Energy Depletion**
  * Highly constrained by the energy source
  * Aimed at depleting the battery power of critical nodes
  II. **Buffer Overflow**
  * Carried out either by filling the routing table with unwanted routing entries or by consuming the data packet buffer space with unwanted data
  * Lead to a large number of data packets being dropped, leading to the loss of critical information
- **Host Impersonation** – A compromised internal node can act as another node and respond with appropriate control packets to create wrong route entries, and can terminate the traffic meant for the intended destination node
- **Information Disclosure** – A compromised node can act as an informer by deliberate disclosure of confidential information to unauthorized nodes
- **Interference** – A common attack in defense applications to jam the wireless communication by creating a wide spectrum noise

Energy Management

• Energy management is defined as the process of managing the sources & consumers of energy in a node (or in the network) for enhancing the lifetime of a network.

• Features of energy management are

→Shaping the energy discharge pattern of a node's battery to enhance battery life
→Finding routes that consumes minimum energy
→Using distributed scheduling schemes to improve battery life
→Handling the processor & interface devices to minimize power consumption

• Energy management can be classified into the following categories

a. Transmission Power Management

- The power consumed by the Radio Frequency (RF) module of a mobile-node is determined by several factors such as

  * State of operation
  * Transmission power &
  * Technology used for the RF circuitry

- The state of operation refers to transmit, receive, and sleep modes of the operation.
- The transmission power is determined by

  → Reachability requirement of the network
  → Routing protocol &
  → MAC protocol employed

b. Battery Energy Management

- The battery management is aimed at extending the battery-life of a node

  → by taking advantage of its chemical properties, discharge patterns and
  → by selection of a battery from a set of batteries that is available

c. Processor Power Management

- The clock-speed and the number of instructions executed per unit time are some of the processor parameters that affect power consumption.
- The CPU can be put into different power saving modes during low processing load conditions.
- The CPU power can be completely turned off if the machine is idle for a long time. In such a case, interrupts can be used to turn on the CPU upon detection of user interaction.

d. Devices Power Management

- Intelligent device management can reduce power consumption of a mobile node significantly.
- This can be done by the OS(operating system)

  → by selectively powering down interface devices that are not used or
  → by putting devices into different power saving modes, depending on their usage

Deployment Considerations

The deployment of a commercial adhoc network has the following benefits when compared to wired networks

a) Low Cost of Deployment

- The use of multi-hop wireless relaying eliminates the requirement of cables & maintenance in deployment of communication-infrastructure.
- The cost involved is much lower than that of wired networks.

b) Incremental Deployment

- Deployment can be performed incrementally over geographical regions of the city.
- The deployed part of the network starts functioning immediately after the minimum configuration is done.

c) Short Deployment Time

- Compared to wired networks, the deployment-time is considerably less due to the absence of any wired links.

d) Reconfigurability

- The cost involved in reconfiguring a wired network covering a MAN is very high compared to that of an adhoc network covering the same service area.

The following are the major issues to be considered in deploying an adhoc network

a) Scenario of Deployment

• The scenario of deployment has significance because the capability required for a mobile-node varies with the environment in which it is used.

• The following are some of the different scenarios

Military Deployment

- It can be either,

  * *Data-centric network:* Handle a different pattern of data traffic & can be partially comprised of static nodes. Eg: wireless sensor network

  * *User-centric network:* Consists of highly mobile nodes with or without any support from any infrastructure. Eg: armored vehicles carrying soldiers equipped with wireless devices

Emergency Operations Deployment

- Demands a quick deployment of rescue personnel equipped with hand-held communication equipment.
- The network should provide support for time-sensitive traffic such as voice & video.
- Short data messaging can be used in case the resource constraints do not permit voice communication.

Commercial Wide-Area Deployment

- The aim of the deployment is to provide an alternate communication infrastructure for wireless communication in urban areas & areas where a traditional cellular base-station cannot handle the traffic volume. Eg: wireless mesh networks

Home Network Deployment

- Deployment needs to consider the limited range of the devices that are to be connected by the network. Eg: short transmission range avoid network partitions

b) Required Longevity of Network

• If the network is required for a short while, battery-powered mobile nodes can be used.

• If the connectivity is required for a longer duration of time, fixed radio relaying equipment with regenerative power sources can be deployed.

c) Area of Coverage

• Determined by the nature of application for which the network is set up.

Eg: home area network is limited to the surroundings of a home

• The mobile nodes' capabilities such as the transmission range & associated hardware, software, & power source should match the area of coverage required.

d) Service Availability

• It is defined as the ability of an adhoc network to provide service even with the failure of certain nodes.

• It has significance

* in a fully mobile adhoc network used for tactical communication &

* in partially fixed adhoc network used in commercial communication infrastructure such as WMNs.

e) Operational Integration with other Infrastructure

• Considered for improving the performance or gathering additional information, or for providing better QoS.

• In military environment, integration of adhoc networks with satellite networks improves the capability of the adhoc networks.

f) Choice of Protocol

• The choice of protocols at different layers of the protocol stack is to be done taking into consideration the deployment scenario.

• A TDMA-based & insecure MAC protocol may not be the best suited compared to a CDMA-based MAC protocol for a military application.

AD HOC WIRELESS INTERNET

- Adhoc wireless internet extends the services of the internet to the end-users over an adhoc network (Fig 5.7).
- Some of the applicationsare

  * Wireless mesh network
  * Provisioning of temporary internet services to major conference venues
  * Sports venues
  * Temporary military settlements
  * Battlefields &
  * Broadband internet services in rural regions

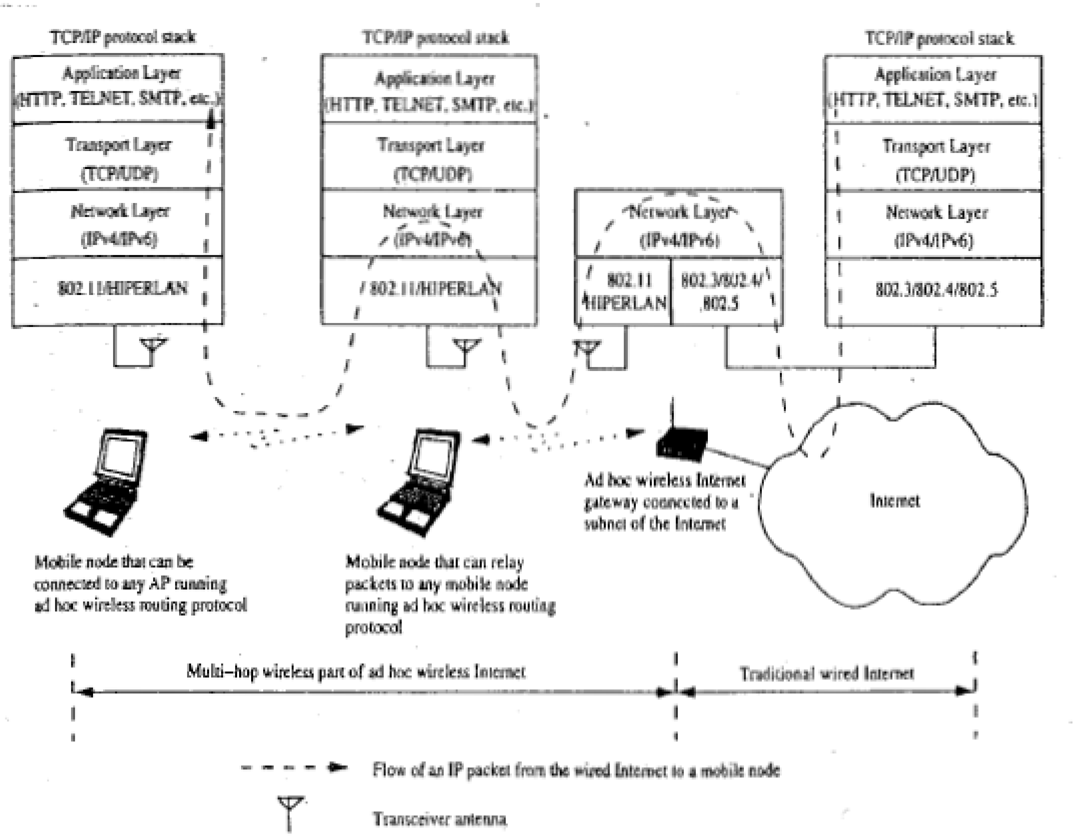

TCP/IP protocol stack
Application Layer (HTTP, TELNET, SMTP, etc.)
Transport Layer (TCP/UDP)
Network Layer (IPv4/IPv6)
802.11/HIPERLAN
TCP/IP protocol stack
Application Layer (HTTP, TELNET, SMTP, etc.)
Transport Layer (TCP/UDP)
Network Layer (IPv4/IPv6)
802.11/HIPERLAN
Network Layer (IPv4/IPv6)
802.11 HIPERLAN
802.3/802.4/802.5
TCP/IP protocol stack
Application Layer (HTTP, TELNET, SMTP, etc.)
Transport Layer (TCP/UDP)
Network Layer (IPv4/IPv6)
802.3/802.4/802.5
Ad hoc wireless Internet gateway connected to a subnet of the Internet
Internet
Mobile node that can be connected to any AP running ad hoc wireless routing protocol
Mobile node that can relay packets to any mobile node running ad hoc wireless routing protocol
Multi-hop wireless part of ad hoc wireless Internet
Traditional wired Internet
Flow of an IP packet from the wired Internet to a mobile node
Transceiver antenna

Figure 5.7 Schematic diagram of adhoc wireless internet

The major issues to be considered for a successful adhoc wireless internet are the following:

1. Gateway

- They are the entry points to the wired-internet.
- Generally, they are owned & operated by a service-provider.
- They perform following tasks

  * Bandwidth management
  * Load balancing
  * Traffic shaping
  * Packet filtering &
  * Address, service & location discovery

2. Address Mobility

- This problem is worse here as the nodes operate over multiple hops.
- Solution such as Mobile IP can provide temporary alternative.

3. Routing

- It is a major problem due to

  * dynamic topological changes
  * presence of gateways
  * multi-hop relaying &
  * hybrid character of network

- Possible solution is to: use separate routing protocol for the wireless part of adhoc wireless internet.

4. Transport Layer Protocol

- Several factors are to be considered here, the major one being the state-maintenance-overhead at the gateway-nodes.

5. Load Balancing

- It is essential to distribute the load so as to avoid the situation where the gateway-nodes become bottleneck-nodes.

6. Pricing/Billing

- Since internet-bandwidth is expensive, it is very important to introduce pricing/billing strategies for the adhoc network.

7. Provisioning of Security

- Security is a prime concern, since the end-users can utilize the adhoc network to make e-commerce transaction.

8. QoS Support

- Provisioning of QoS-support is a very important issue because of
  * widespread use of voice over IP(VOIP) &
  * growing multimedia applications over the internet

9. Service, Address & Location Discovery

- Service discovery refers to the activity of identifying the party which provides the service(or resource).
- Address discovery refers to the services such as those provided by ARP or DNS operating within the wireless domain.
- Location discovery refers to different activities such as
  * detecting location of a particular mobile-node in network or
  * detecting geographical location of nodes

UNIT 2: MAC – 1

ISSUES IN DESIGNING MAC PROTOCOL FOR ADHOC WIRELESS NETWORK

Bandwidth Efficiency

- It is defined as the ratio of the bandwidth utilized for data transmission to the total available bandwidth.
- Bandwidth must be utilized in efficient manner.
- Control overhead must be kept as minimal as possible.

Quality of Service support

- This is essential for supporting time-critical traffic sessions.
- The protocol should have resource reservation mechanism that takes into considerations

  → the nature of wireless channel and

  → the mobility of nodes

Synchronization

- This is very important for bandwidth (time slot) reservation by nodes.
- The protocol must consider synchronization between nodes in the network.
- Exchange of control packets may be required for achieving time synchronization among nodes.

Hidden and Exposed Terminal Problems

- The hidden terminal problem refers to the collision of packets at a receiving node due to the simultaneous transmission of those nodes that are not within the direct transmission range of the sender but are within the transmission range of the receiver.
- Collision occurs when both nodes transmit packets at the same time without knowing about the transmission of each other.

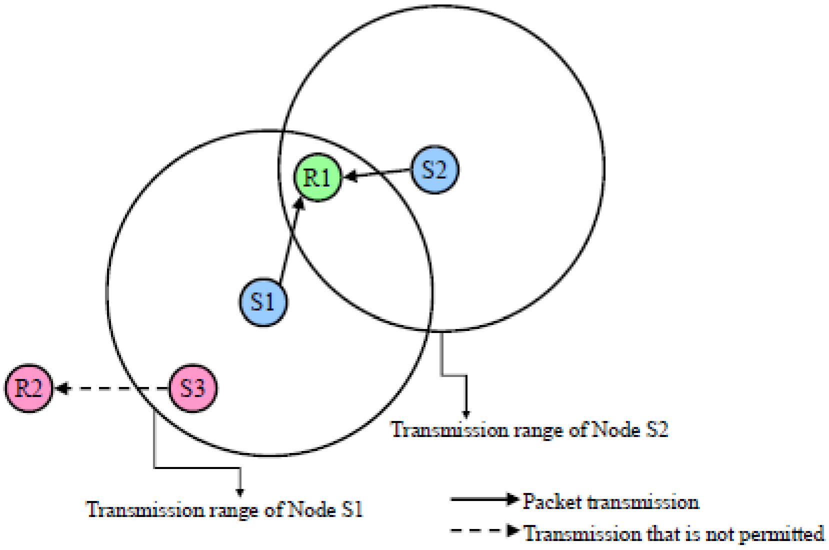

R1
S2
S1
R2
S3
Transmission range of Node S2
Transmission range of Node S1
Packet transmission
Transmission that is not permitted

Figure 6.1 Hidden and exposed terminal problems

- In figure 6.1, S1 and S2 are hidden from each other & they transmit simultaneously to R1 which leads to collision.
- The exposed terminal problem refers to the inability of a node, which is blocked due to transmission by a nearby transmitting node, to transmit to another node.
- If S1 is already transmitting to R1, then S3 cannot interfere with on-going transmission & it cannot transmit to R2.
- Hidden & exposed terminal problems reduce the throughput of a network when traffic load is high.

Error-prone Shared Broadcast Channel

- When a node is receiving data, no other node in its neighborhood (apart from the sender) should transmit.
- A node should get access to the shared medium only when its transmission do not affect any ongoing session.
- The protocol should grant channel access to nodes in such a manner that collisions are minimized.
- Protocol should ensure fair bandwidth allocation.

Distributed Nature

- There is no central point of coordination due to the mobility of the nodes.
- Nodes must be scheduled in a distributed fashion for gaining access to the channel.

Mobility of Nodes

- Nodes are mobile most of the time
- The protocol design must take this mobility factor into consideration so that the performance of the system is not affected due to node mobility.

DESIGN GOALS OF A MAC PROTOCOL FOR AD HOC WIRELESS NETWORKS

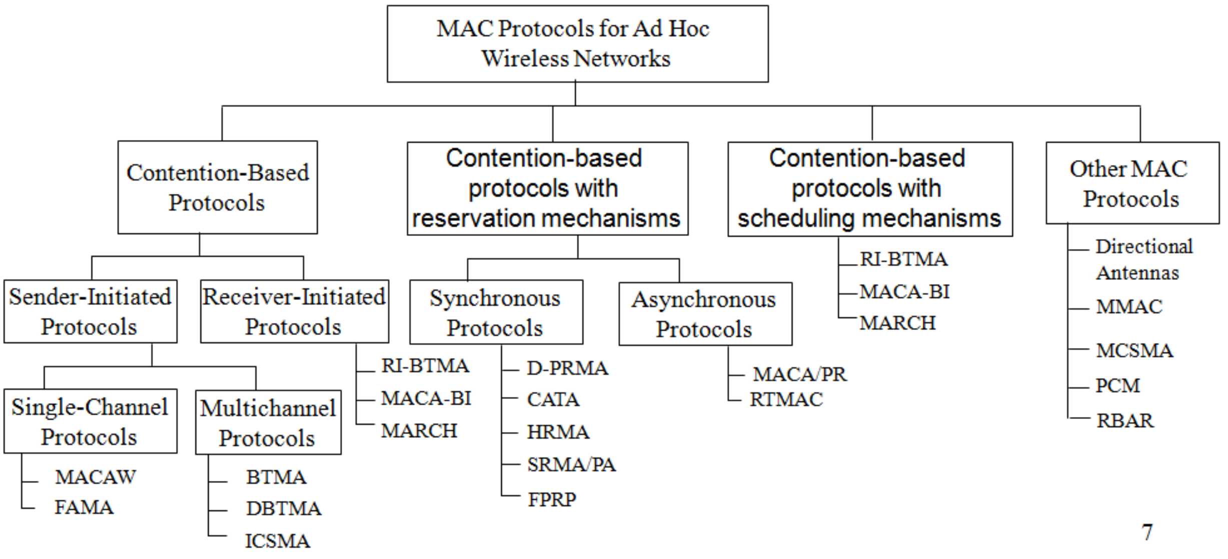
- The available bandwidth must be utilized efficiently.
- Control overhead must be kept as low as possible.
- The operation of a protocol should be distributed.
- The access delay must be kept low. (Access delay refers to the average delay experienced by any packet to get transmitted).
- The protocol should provide QoS support for real-time traffic.
- The protocol should minimize the effects of hidden and exposed terminal problems.
- The protocol should provide time synchronization among nodes.
- The protocol should ensure fair allocation of bandwidth to nodes.
- The protocol must be scalable to large networks.
- The protocol should have power control mechanisms in order to efficiently manage energy consumption of the nodes.
- The protocol should have mechanisms for adaptive data-rate control.
- The protocol should try to use directional antennas which can provide advantages such as

  → reduced interference
  → increased spectrum reuse &
  → reduced power consumption

CLASSIFICATION OF MAC PROTOCOLS

Ad hoc network MAC protocols can be classified into three basic types:

i. Contention-based protocols
ii. Contention-based protocols with reservation mechanisms
iii. Contention-based protocols with scheduling mechanisms
iv. Other MAC protocols [protocols which do not fall under above 3 categories]



Contention-Based Protocols

- Here, the channel access policy is based on competition.
- Whenever a node needs to send a packet, it tries to get access to the channel.
- These protocols cannot provide QoS, since access to the network cannot be guaranteed beforehand.
- It can be further subdivided into:

→*Sender-initiated protocols:* Packet transmissions are initiated by the sender node.

* Single-channel sender-initiated protocols: A node that wins the contention to the channel can make use of the entire bandwidth.

* Multichannel sender-initiated protocols: The available bandwidth is divided into multiple channels. Many nodes can simultaneously perform data transmission using multiple channels.

→*Receiver-initiated protocols:* The receiver node initiates the contention resolution protocol.

Contention-Based Protocols with Reservation Mechanisms

- These protocols provide bandwidth reservation ahead; therefore, they can provide QoS support.
- These can be further subdivided into:

→ *Synchronous protocols*: There is time synchronization among all nodes in the network; the nodes in the neighborhood are informed of the reservations;

→*Asynchronous protocols:* no global synchronization is needed. Relative time is used for the reservations.

Contention-Based Protocols with Scheduling Mechanisms

- There can be

-packet scheduling at the nodes or
-node scheduling for access to the channel

- Node scheduling should not treat the nodes unfairly.
- Some of these protocols consider battery power in their node scheduling.

Other protocols are those MAC protocols that do not strictly fall under the above categories.

MACA

- When a node wants to transmit a data packet, it first transmits a RTS (Request To Send) frame.
- The receiver node, on receiving the RTS packet, if it is ready to receive the data packet, transmits a CTS (Clear to Send) packet.
- Once the sender receives the CTS packet without any error, it starts transmitting the data packet.
- If a packet transmitted by a node is lost, the node uses the binary exponential back-off (BEB) algorithm to back off a random interval of time before retrying.
- The binary exponential back-off mechanism used in MACA might starve flows sometimes. The problem is solved by MACAW.

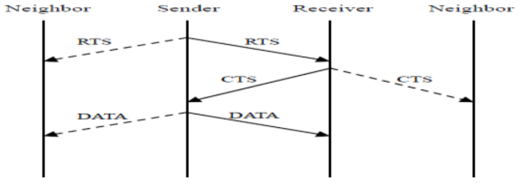


Figure 6.3 Packet transmission in MACA

CONTENTION BASED PROTOCOLS

MACAW (MACA for Wireless)

- Back-off mechanism used in MACA starves flows
- To prevent large variations in the back-off values, a multiplicative increase and linear decrease (MILD) is used in MACAW
  - →On Collision: back-off is increased by a multiplicative factor (1.5)
  - →On Successful transmission: back-off is decreased by one
- The sender senses the carrier to see and transmits a RTS (Request To Send) frame if no nearby station transmits a RTS.
- The receiver replies with a CTS (Clear To Send) frame.
- Sender sends DATA, for which receiver responds with ACK.
- RTS/CTS packets carry the expected duration of the data transmission.

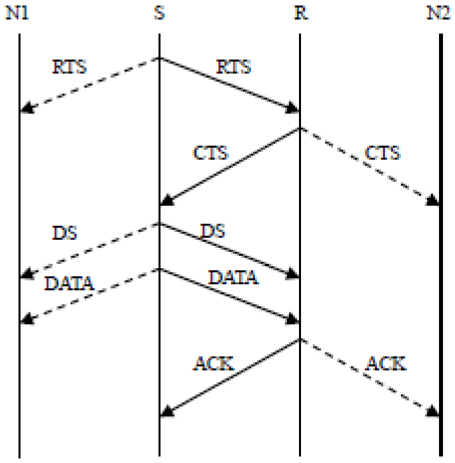
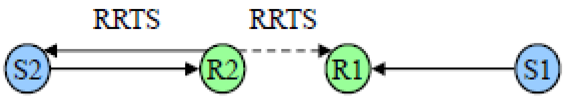


Figure 6.7 Packet exchange in MACAW.

- A node near the receiver on hearing the CTS packet, defers its transmission till the receiver receives the data packet. This overcomes hidden node problem.
- A node near the sender that only hears the RTS is free to transmit simultaneously when the sender is transmitting data. This overcomes exposed node problem.
- The receiver sends an ACK when receiving a frame.
  - → Neighbors keep silent until see ACK.
- Collision handling: If a packet is lost (collision), the node uses the binary exponential back-off (BEB) algorithm to back off for a random time interval before retrying.
- RTS/CTS mechanism does not solve the exposed terminal problem.
  - → Solution: New control packet called data-sending (DS) can be used. DS contains information such as the duration of the forthcoming data transmission.
- The protocol uses one more control packet called the request-for-request-to-send (RRTS)
- Synchronization information needs to be propagated to the concerned nodes
- If a node had received an RTS previously for which it was not able to respond because there exists on-going transmission, then it waits for the next contention period and transmits RRTS



Floor Acquisition Multiple Access Protocols (FAMA)

- It is based on a channel access discipline which consists of
  →a carrier-sensing operation and
  →a collision-avoidance dialog between the sender and the intended receiver of a packet
- Floor acquisition refers to the process of gaining control of the channel.
- At any time, only one node is assigned to use the channel.
- Carrier-sensing by the sender,
  →followed by the RTS-CTS control packet exchange,
  →enables the protocol to perform as efficiently as MACA
- Data transmission to be collision free, the duration of an RTS must be at least twice the maximum channel propagation delay
- Two variations of FAMA
  →RTS-CTS exchange with no carrier-sensing uses the ALOHA protocol for transmitting RTS packets (MACA).
  →RTS-CTS exchange with non-persistent carrier-sensing uses non-persistent CSMA for the same purpose (FAMA-NTR).

FAMA-NTR

- Before sending a packet, the sender senses the channel.
- If channel is busy, the sender back-off a random time and retries later.
- If the channel is free, the sender sends RTS and waits for a CTS packet.
- If the sender cannot receive a CTS, it takes a random back-off and retries later.
- If the sender receives a CTS, it can start transmission data packet.
- In order to allow the sender to send a burst of packets, the receiver is made to wait a time duration τ seconds after a packet is received

Busy Tone Multiple Access Protocols (BTMA)

- The transmission channel is split into two:
  → a data channel for data packet transmissions
  → a control channel used to transmit the busy tone signal
- When a node is ready for transmission, it senses the channel to check whether the busy tone is active.
- If not, it turns on the busy tone signal and starts data transmissions.
- Otherwise, it reschedules the packet for transmission after some random rescheduling delay.
- When a node is transmitting, no other node in the two-hop neighborhood of the transmitting node is permitted to simultaneously transmit.
- Drawback: very poor bandwidth utilization.

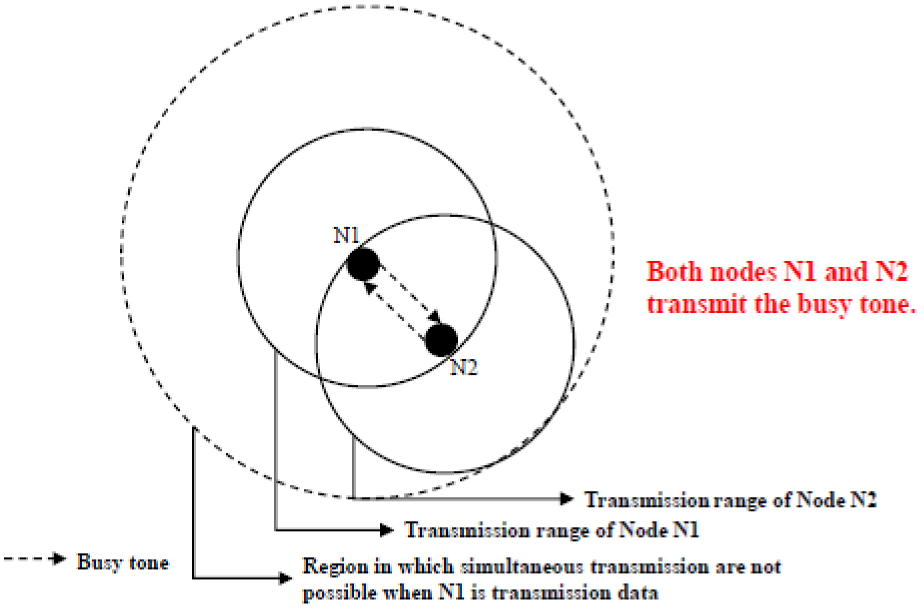

N1
N2
Both nodes N1 and N2 transmit the busy tone.
Transmission range of Node N2
Transmission range of Node N1
Busy tone
Region in which simultaneous transmission are not possible when N1 is transmission data

Figure 6.8 Transmission in BTMA.

Dual Busy Tone Multiple Access Protocol (DBTMAP)

• The transmission channel is divided into:
  → The data channel is used for data packet transmission
  → The control channel is used for RTS, CTS, busy tones

• Use two busy tones on the control channel, BT_t and BT_r.
  → BT_t: indicate that it is transmitting on the data channel
  → BT_r: indicate that it is receiving on the data channel

• Two busy tone signals are two sine waves at different Frequencies

• When a node is ready to transmit a data packet (See Figure 6.9)
  →First, it senses the channel to determine whether the BT_r signal is active
  →If there is no BT_r signal, then it transmit RTS packet
  →On receiving the RTS packets, receiver checks whether the BT_t tone is active
  →If there is no BT_t signal, Receiver Sends CTS packet and turns on the BT_r signal
  →Sender receives CTS, turns on BT_t signal, starts data transmission and turns off BT_t signal
  →Receiver receives data and turn off BT_r signal

• DBTMA has better network utilization than RTS/CTS based protocol

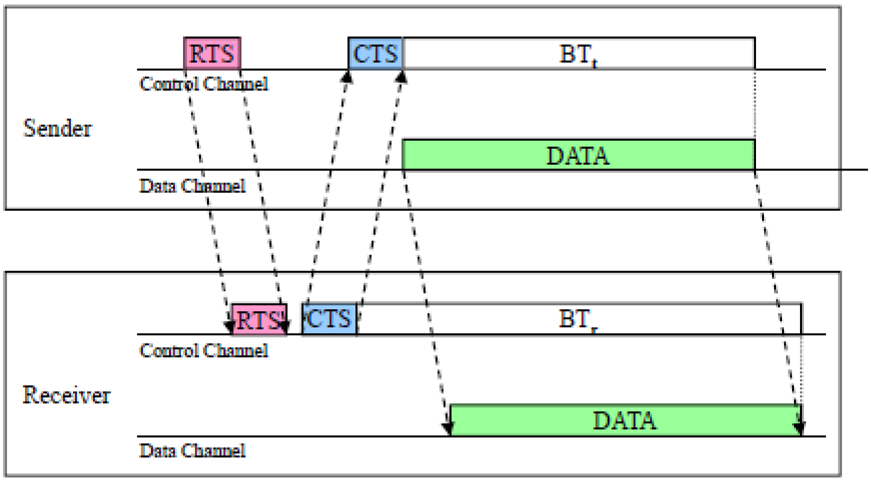


Figure 6.9 Packet transmission in DBTMA.

Receiver-Initiated Busy Tone Multiple Access Protocol (RI-BTMA)

• The transmission channel is split into two:

→a data channel for data packet transmissions

→a control channel used for transmitting the busy tone signal

• A node can transmit on the data channel only if it finds the busy tone to be absent on the control channel.

• The data packet is divided into two portions: a preamble and the actual data packet.

• The busy tone serves two purposes:

-Acknowledges the sender the successful of receiving preamble

-Inform the nearby hidden nodes the impending transmission

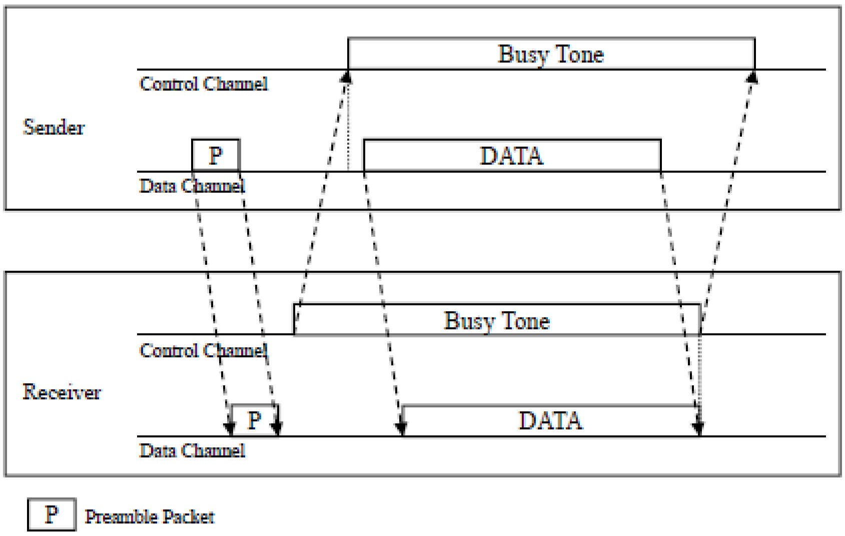


Figure 6.10 Packet transmission in RI-BTMA.

• The operation of the RI-BTMA protocol (See Figure 6.10) two types

→The basic protocol

No backlog buffers: packets that suffer collisions cannot be retransmitted

→The controlled protocol

Backlogged mode: backlog buffer is non-empty

Backlog buffers: transmitting a backlogged packet in next idle slot with a probability *q*

Non-backlogged mode: transmitting a non-backlogged packet in the next idle slot with a probability *p*

MACA-By Invitation Protocol (MAC BI)

• It is a receiver-initiated protocol.

• It reduces the number of control packets used in the MACA protocol.

• It eliminated the need for the RTS packet.

• The receiver node initiates data transmission by transmitting a ready-to-receive(RTR) control packet to the sender.

• If it is ready to transmit, the sender node responds by sending a DATA packet. Thus, data transmission occurs through a two-way handshake mechanism.

• The efficiency of the MACA-BI scheme is mainly dependent on

→the ability of the receiver node to predict accurately the arrival rates of traffic at the sender nodes.

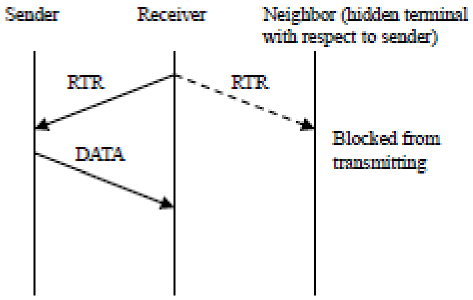

Sender
Receiver
Neighbor (hidden terminal with respect to sender)
RTR
RTR
DATA
Blocked from transmitting

Figure 6.11 Packet transmission in MACA-BI

Media Access with Reduced Handshake Protocol (MARCH)

- It is a receiver-initiated protocol.
- It doesn't require any traffic prediction mechanism.
- It exploits the broadcast nature of traffic from omni-directional antennas to reduce the number of handshakes involved in the data transmission.
- A node obtains information about the data packet arrivals at its neighbouring nodes by overhearing the CTS packets transmitted by them.
- It then sends a CTS packet to the concerned neighbour node for relaying data from that node.
- The throughput of MARCH is significantly high compared to MACA.
- Control overhead is much less.
- Less BW is consumed for control traffic.

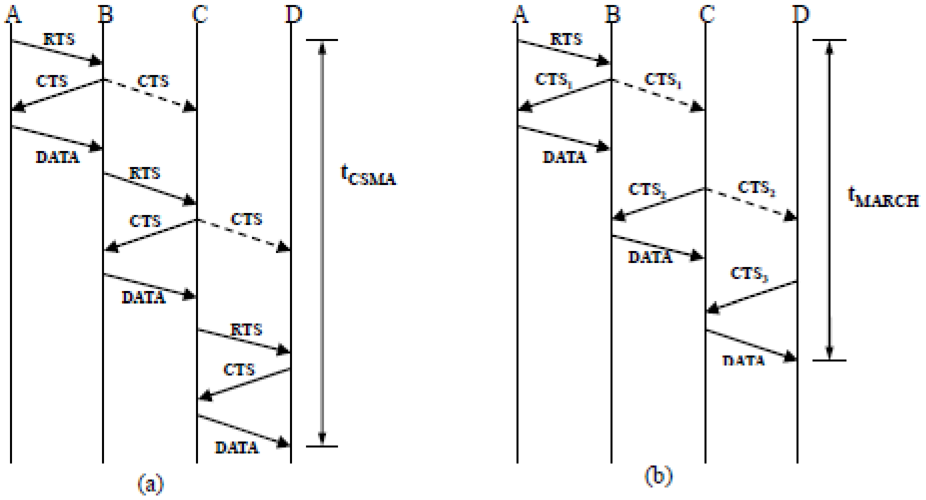


Figure 6.13 Handshake mechanism in (a) MACA and (b) MARCH

CONTENTION BASED PROTOCOLS WITH RESERVATION MECHANISMS

Distributed Packet Reservation Multiple Access Protocol (D-PRMA)

• It is based on TDMA.
• The time division of the channel is done into frames, then further into slots, then further into minislots.
• Each minislot contains two control fields,

RTS/BI – Request To Send / Busy Indication and
CTS/BI – Request To Send / Busy Indication.

• These control fields are used

→for slot reservation and
→for overcoming the hidden terminal problem

• The mechanism of competition for slots is such that a certain period at the beginning of every slot is reserved for carrier-sensing.
• The nodes compete for the first minislot in each slot.
•The winning node transmits a RTS packet through the RTS/BI part of the first minislot.
•The receiver responds by sending a CTS packet through the CTS/BI field. Thus, the node is granted all the subsequent minislots.
• Also, the same slot in each subsequent frame can be reserved for this winning node until it completes its packet transmission session

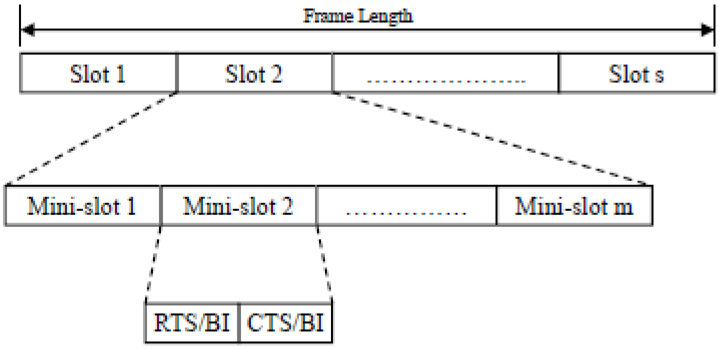


Figure 6.15 Frame structure in D-PRMA

• Within a reserved slot, communication between the source and receiver nodes takes by means of either

→time division duplexing (TDD) or
→frequency division duplexing (FDD)

• There are two rules to prioritize voice terminals over data terminals:

→ 1st rule: Voice terminals starts contenting from mini-slot 1 with probability p = 1 while data terminals can start such content with p < 1
→ 2nd rule: Only the winner of a voice terminal can reserve the same slot in each subsequent frame until the end of packet transmission while the winner of a data terminal can only use one slot

• In order to avoid the hidden terminal problem,

all nodes hearing the CTS sent by the receiver are not allowed to transmit during the remaining period of that same slot

• In order to avoid the exposed terminal problem,

a node hearing the RTS but not the CTS is still allowed to transmit

Collision Avoidance Time Allocation Protocol (CATA)

- It is based on dynamic topology-dependent transmission scheduling.
- Nodes contend for and reserve time slots by means of a distributed reservation and handshake mechanism.
- It supports broadcast, unicast, and multicast transmissions at the same time.
- The operation is based on two basic principles:

  → The receiver of a flow must inform other potential source nodes about the reservation of the slot, and also inform them about interferences in the slot.

  → Negative acknowledgements are used at the beginning of each slot for distributing slot reservation information to senders of multicast sessions.
- Time is divided into frames, each frame into slots, and each slot into 5 minislots.
- The first 4 minislots are used for transmitting control packets and are called control minislots (CMS).

  While the last minislot is used for data transmission and is called data minislot (DMS).

  CMS1: Slot Reservation (SR)
  CMS2: RTS
  CMS3: CTS
  CMS4: not to send (NTS)
  DMS: Data transmission

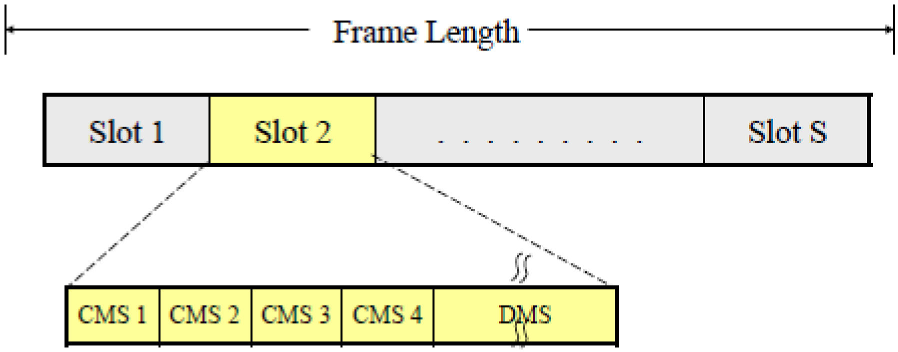


Figure 6.16. Frame format in CATA.

- The CMS1 and CMS2 are used to inform neighbors about the current reservation.

  While CMS3 and CMS4 are used for channel reservation.

•Each node that receives data during the DMS of current slot transmits an SR in CMS1. This serves to inform other neighbouring potential sender nodes about the currently active reservations

- Every node that transmits data during the DMS of current slot transmits an RTS in CMS2
- CMS3 and CMS4 are used as follows:

  → The sender of an intend reservation, if it senses the channel is idle in CMS1, transmits an RTS in CMS2

  → Then the receiver transmits a CTS in CMS3

  → If the reservation was successful, the data can be transmitted in current slot and the same slot in subsequent frames

  → Once the reservation was successfully, in the next slot both the sender and receiver do not transmit anything during CMS3

  → During CMS4, the sender transmits a NTS(NTS serves as a negative acknowledgement)

  →A potential multicast source node that receives the NTS packet understands that its reservation is failed
- It works well with simple single-channel half-duplex radios
- It is simple and provides support for collision-free broadcast and multicast traffic

Hop Reservation Multiple Access Protocol (HRMA)

- It is a time slot-reservation protocol where each slot is assigned a separate frequency channel.
- A handshake mechanism is used for reservation to enable node pairs to reserve a frequency hop, thus providing collision-free communication and avoiding the hidden terminal problem

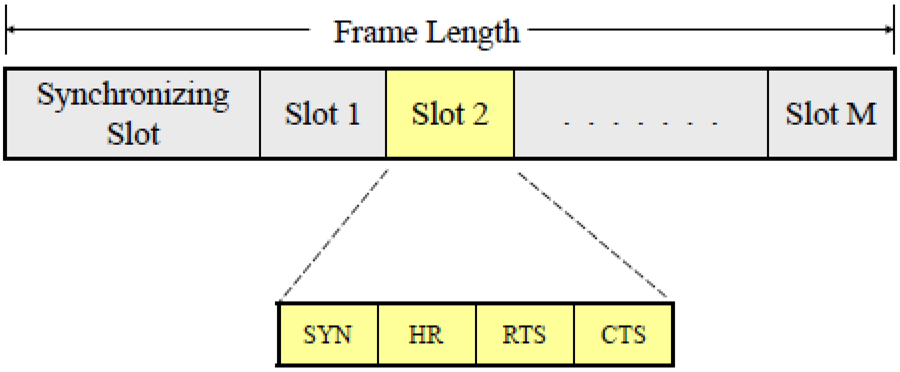


Figure 6.17. Frame format in HRMA .

- One frequency channel is a dedicated synchronizing channel where nodes exchange synchronization information.
  The remaining frequency channels are paired,
  →one channel in each pair is used for hop-reservation packets(RTS & CTS) & data packets
  →the other one is used for acknowledgement (ACK)
- Time is slotted and each slot is assigned a separate frequency hop.
- Each time slot is divided into four periods, namely, synchronising period, HR period, RTS period, and CTS period.
- Each period meant for transmitting or receiving the synchronising packet, FR packet, RTS packet, and CTS packet respectively.
- After the handshaking is over, the two nodes communicate by sending data and ACKs on the very same frequency channels.
- All idle nodes hop to the synchronizing frequency *f0* and exchange synchronization information.
- Synchronizing slot: is used to identify
  →the beginning of a frequency hop and
  →the frequency to be used in the immediately following hop
- A node ready to transmit data,
  →it senses the HR period of the current slot
  →If the channel is idle during HR period; it transmits an RTS during RTS period and waits for CTS during CTS period
  → On receiving the RTS, the destination node transmits the CTS packet during the CTS period of the same slot and waits for the data packet
  → If the source node receives the CTS packet correctly, it implies that the source and receiver nodes have successfully reserved the current hop
  → If the channel is busy during HR period, it backs off for a randomly multiple slots
- Suppose the sender needs to transmits data across multiple frames, it informs the receiver through the header of the data packet
  →The receiver node transmits an HR packet during the HR period of the same slot in next frame to informs its neighbors
  →The sender receiving the HR packet, it sends an RTS during the RTS period and jams other RTS packets
  →Both receiver remain silent during the CTS period

Soft Reservation Multiple Access with Priority Assignment (SRMA/PA)

• It is developed with the main objective of supporting integrated services of real-time and non-real-time application in ad hoc networks.

• Nodes use

→a collision-avoidance handshake mechanism and

→a soft reservation mechanism

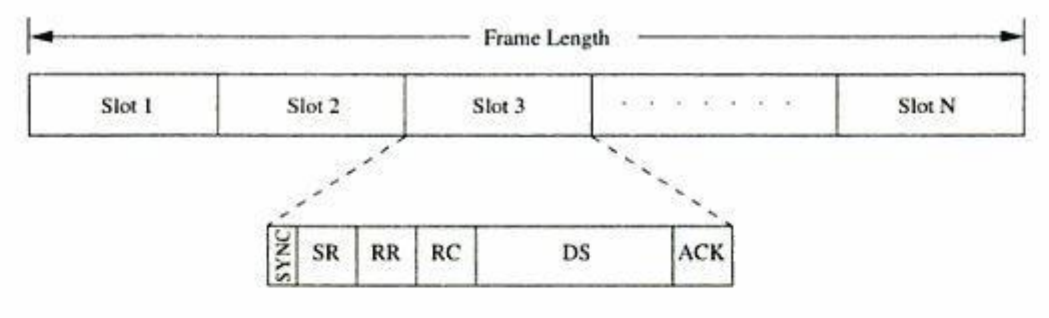


Figure 6.19. Frame structure in SRMA/PA.

• Time is divided into frames, with each frame consisting of a fixed number of slots.

• Each slot is further divided into 6 different fields namely SYNC, soft reservation (SR), reservation request (RR), reservation confirm (RC), data sending (DS) and acknowledgement (ACK).

→The SYNC field is used for synchronization purposes

→The SR, RR, RC, & ACK fields are used for transmitting & receiving the corresponding control packets

→The DS field is used for data transmission

→The SR packet serves as a busy tone. It informs the nodes about the reservation of the slot. It also carries the access priority value assigned to the node that has reserved the slot

• A node determines whether or not a slot is free through the SR field of that slot.

• When an idle node receives a data packet for transmission, the node waits for a free slot and transmits the RR packet in the RR field of that slot.

• In case of a voice terminal node, the node tries to take control of the slot already reserved by a data terminal if it finds it priority level to be higher than that of the data terminal. This process is called soft reservation.

• Priority levels are initially assigned to nodes based on the service classes in a static manner.

• It is required that priority of voice terminal pv(R) > priority of data terminal pd(R).

• A node can be in one of the two states:

→ A node is said to be in the active state if it is currently transmitting

→A node is said to be in the idle state if it does not have any packet to be transmitted

• In the active state itself, nodes can be in one of the two states: access state and reserved state.

• Access state is one in which the node is backlogged and is trying to reserve a slot for transmission.

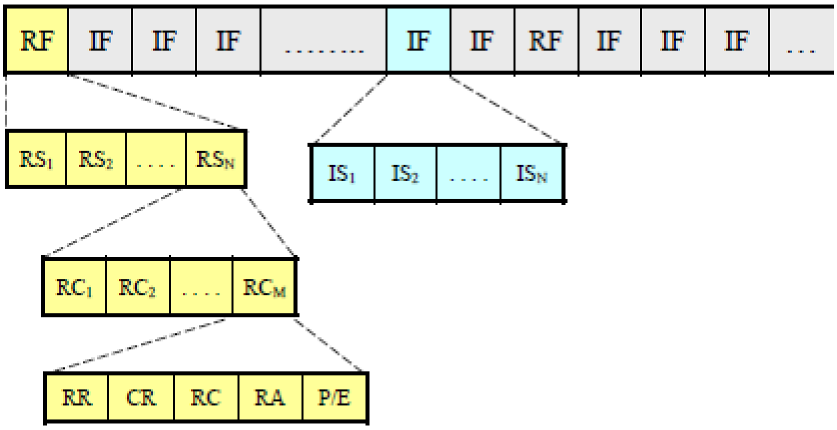
• In order to avoid collisions,

→a binary exponential back-off algorithm is used for non-real time connections and

→a modified binary exponential back-off algorithm is used for real time connection

Five-Phase Reservation Protocol (FPRP)

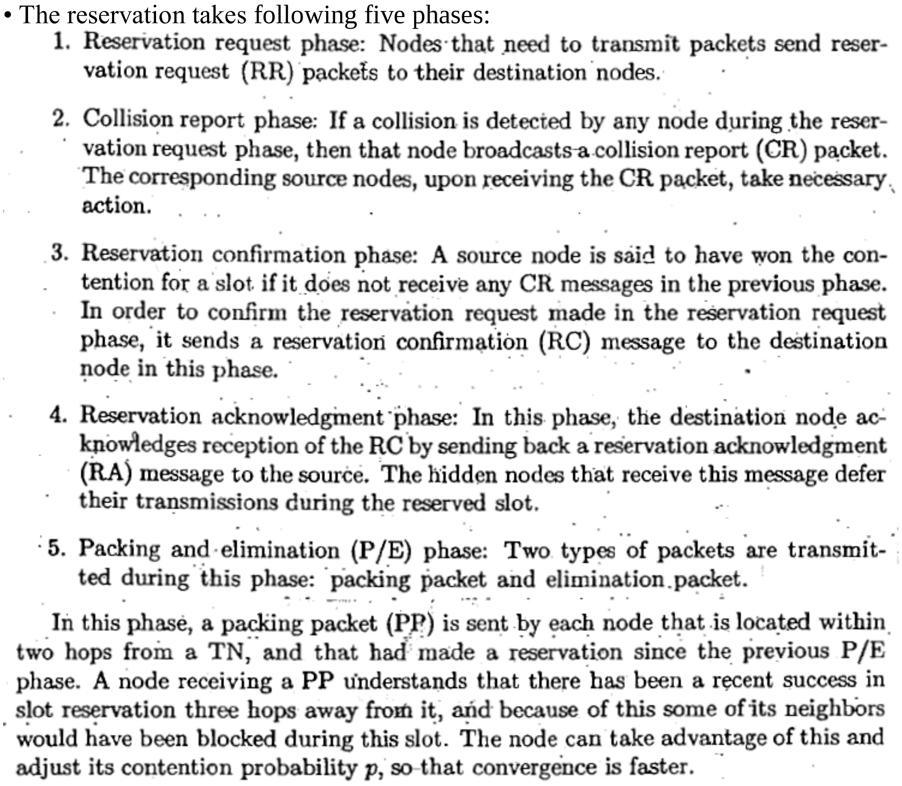
- It is a single-channel TDMA-based broadcast scheduling protocol.
- The protocol is fully distributed, that is, multiple reservations can be simultaneously made throughout the network.
- The protocol assumes the availability of global time at all nodes.
- No ordering among nodes is followed
- Nodes need not wait for making time slot reservations



Five-phase reservation dialog

Figure 6.21. Frame structure in FPRP.

- Time is divided into frames: reservation frame (RF) and information frame (IF).
- Each RF has N reservation slots (RS) and each IF has N information slots (IS).
- Each RS is composed of M reservation cycles (RCs).
- Each RF is followed by a sequence of Ifs.
- In order to reserve an IS, a node needs to contend during the corresponding RS.
- Based on these contentions, a TDMA schedule is generated in the RF and is used in the subsequent Ifs until the next RF.
- During the corresponding IS, a node would be in one of the three states: transmit(T), receive(R) or blocked(B)
- The reservation takes following five phases:



1. Reservation request phase: Nodes that need to transmit packets send reservation request (RR) packets to their destination nodes.

2. Collision report phase: If a collision is detected by any node during the reservation request phase, then that node broadcasts a collision report (CR) packet. The corresponding source nodes, upon receiving the CR packet, take necessary action.

3. Reservation confirmation phase: A source node is said to have won the contention for a slot if it does not receive any CR messages in the previous phase. In order to confirm the reservation request made in the reservation request phase, it sends a reservation confirmation (RC) message to the destination node in this phase.

4. Reservation acknowledgment phase: In this phase, the destination node acknowledges reception of the RC by sending back a reservation acknowledgment (RA) message to the source. The hidden nodes that receive this message defer their transmissions during the reserved slot.

5. Packing and elimination (P/E) phase: Two types of packets are transmitted during this phase: packing packet and elimination packet.

In this phase, a packing packet (PP) is sent by each node that is located within two hops from a TN, and that had made a reservation since the previous P/E phase. A node receiving a PP understands that there has been a recent success in slot reservation three hops away from it, and because of this some of its neighbors would have been blocked during this slot. The node can take advantage of this and adjust its contention probability p, so that convergence is faster.

MACA with Piggy-Backed Reservation (MACA/PR)

• It is based on the MACAW protocol with non-persistent CSMA

• The main components are:

→A MAC protocol
→A reservation protocol
→A QoS routing protocol

• It differentiates real-time packets from the best-effort packets

• It provides guaranteed bandwidth support for real-time packets

Also, it provides reliable transmission of best efforts packets.

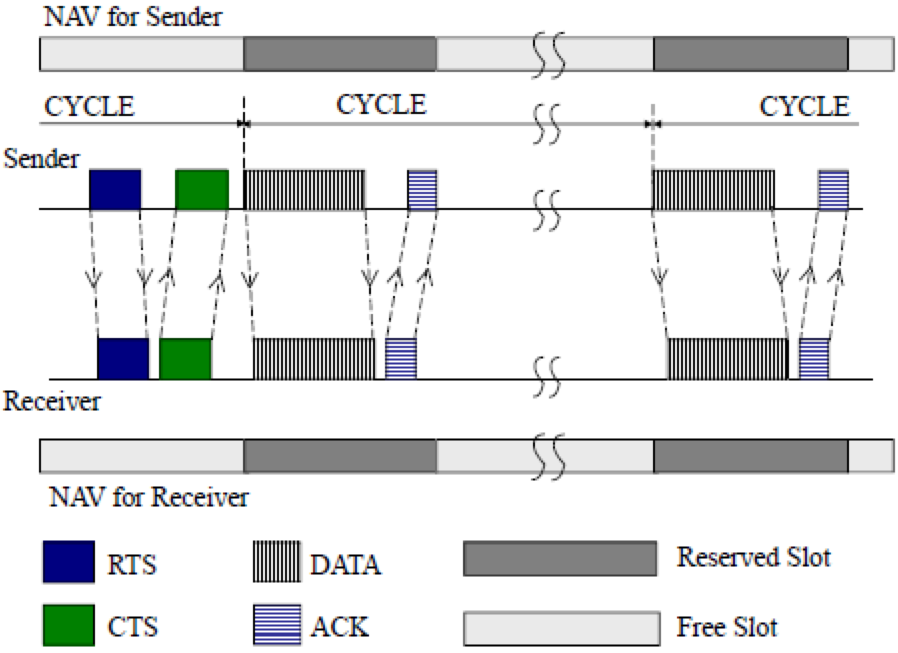


Figure 6.23. Packet transmission in MACA / PR.

• Time is divided into slots.

• Each node records the transmit and receive reservations of its neighbors in a reservation-table(RT).

• For real-time traffic:

→The source first sends an RTS packet, for which the receiver responds with a CTS packet
→Now the source sends the first DATA packet of the real-time session
→Reservation information for the next DATA packet is piggy-backed on this current DATA packet.
→On receiving this DATA packet, the receiver updates its reservation table with the piggy-backed reservation information
→The receiver then sends ACK packet back to the source
→Receiver piggy-backs the reservation confirmation information on the ACK packet

• Advantage: It does not require global synchronization among nodes

Drawback: A free slot can be reserved only if it can fit the entire RTS-CTS-DATA-ACK exchange

• For non-real-time traffic:

→A node that wants to transmit a non-real-time packet, finds a free slot in the table
→Then, it waits for the same slot the next time around
→If it is still free, it sends a RTS packet in the slot, expects a CTS packet, then sends the data and receives the acknowledgement still in the same slot
→The RTS and CTS packets contains the amount of time that the data transmission is going to take place
→In this way, the neighbors of the source and destination nodes can update their tables

Real-Time Medium Access Control Protocol (RTMAC)

- It provides a bandwidth reservation mechanism for supporting real-time traffic.
- It has two components:

1) QoS routing protocol is responsible for
→end-to-end reservation &
→release of bandwidth resources

2) MAC protocol is responsible for
→medium access for best effort traffic &
→reservation for real time traffic

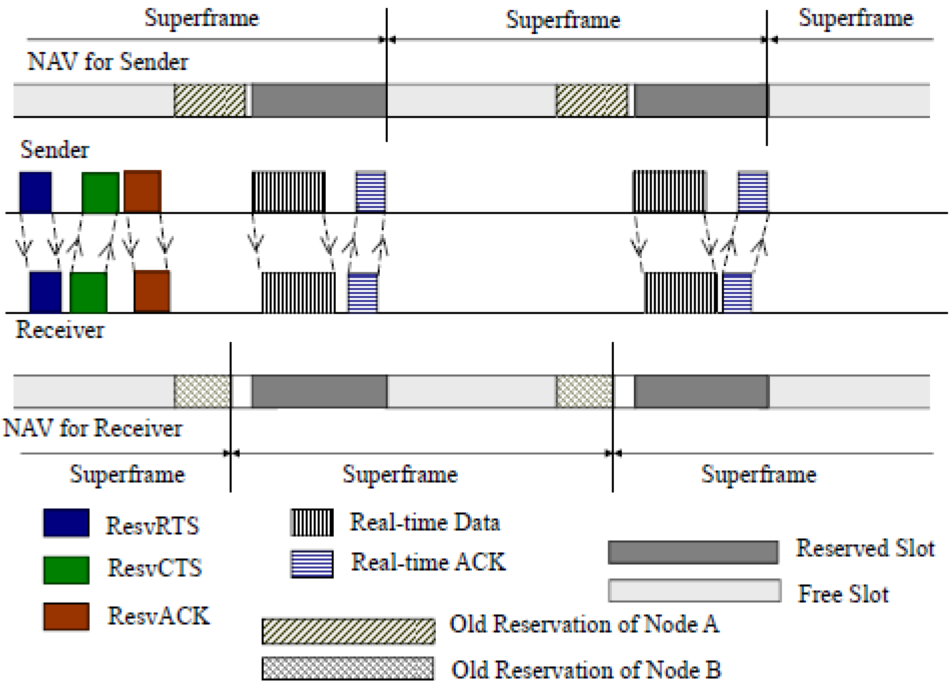


Figure 6.24. Reservation mechanism in RTMAC.

- For transmitting best effort packets: RTS, CTS, and ACK are used
For transmitting real time packets: ResvRTS, ResvCTS, and ResvACK are used
- Time is divided into superframes (Figure 6.24).
- Each superframe consists of a number of reservation-slots (resv).
- A node that needs to transmit real-time packets, first reserves a set of resv-slots.
- The set of resv-slots reserved by a node for a connection on a superframe is called a *connection-slot.*
- The duration of each resv-slot is twice the maximum propagation delay.
- Each node maintains a reservation table (RT).
- RT contains information such as
→ sender-id & receiver-id
→starting and ending times of active reservation
- NAV indicates the network allocation vector maintained at each node.
- Main advantages:
→ Bandwidth efficiency
→ Asynchronous mode of operation where nodes do not require any global time synchronization
→ Flexibility of slot placement in the superframe

UNIT 3: MAC – 2

CONTENTION BASED MAC PROTOCOLS WITH SCHEDULING MECHANISMS

DISTRIBUTED PRIORITY SCHEDULING (DPS)

- It uses the basic RTS-CTS-DATA-ACK packet exchange mechanism (Figure 6.25).
- The protocol works as follows:

  i) When source transmits a RTS, priority-tag of current DATA is piggy-backed on RTS

  ii) On receiving RTS, the receiver responds with CTS.

  iii) The receiver copies priority-tag from the received-RTS and piggy-backs it along

  iv) Neighbors

  → receive the RTS or CTS

  → retrieve the piggy-backed information and

  → make a corresponding entry in their scheduling-tables (STs contain information about packets, which were originally piggy-backed on control and data packets).

  1) When source transmits a DATA, its head-of-line(HOL) packet information is piggy-backed on DATA (HOL packet of a node refers to the packet to be transmitted next by the node).

  2) On receiving DATA, the receiver responds with ACK.

  3) The receiver copies the HOL-information from the received-DATA and piggy-backs it along

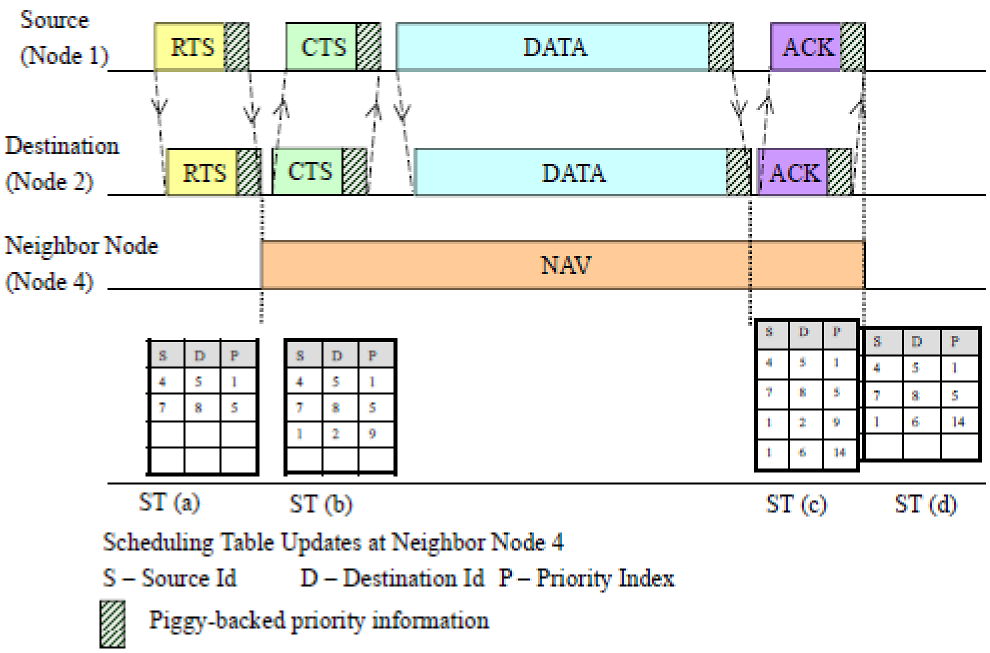
  4) Neighbors

  → receive the DATA or ACK

  → retrieve the piggy-backed information and

  → make a corresponding entry in their STs

  5) When a node hears an ACK, it removes from its ST any entry made earlier for corresponding DATA.



S	D	P
4	5	1
7	8	5

S	D	P
4	5	1
7	8	5
1	2	9

S	D	P
4	5	1
7	8	5
1	2	9
1	6	14

S	D	P
4	5	1
7	8	5
1	6	14

Figure 6.25. Piggy-backing and scheduling table update mechanism in DPS.

MULTI-HOP COORDINATION

- The excess delay incurred by a packet at the upstream-nodes is compensated for at the downstream-nodes.
- When a node receives a packet, it would have already received the priority-index of the packet piggy-backed on the previous RTS packet.
- In case the node is an intermediate-node (which has to further forward the packet), the node calculates the new priority-index of the DATA packet based on the received value of the priority-index.
- If a packet suffers due to excess delay at the upstream-nodes,

  then the downstream-nodes increase priority of packet so that packet is able to meet its end-to-end delay target
- Similarly, if a packet arrives very early due to lack of contention at the upstream-nodes,

  then the priority of that packet would be reduced at the downstream-nodes

DISTRIBUTED WIRELESS ORDERING PROTOCOL (DWOP)

- Packets access the medium according to order specified by an ideal reference scheduler such as FIFO (or earliest deadline first).
- In FIFO, packet priority-indices are set to the arrival-times of packets.
- Each node builds up a scheduling-table(ST) ordered according to the overheard arrival-times.
- It may not suffer due to information asymmetry (Figure 6.26). (Since in most networks, all nodes are not within the radio range of each other, a transmitting node might not be aware of the arrival times of packets queued at another node which is not within its direct transmission range).
- Control packets (RTS/CTS) are used to piggy-back priority-information regarding HOL-packets of nodes.
- Key concept: A node is made eligible to contend for the channel only if its locally queued packet has a smaller arrival-time compared to all other arrival-times in its ST.
- Following two additional table management techniques are used in order to keep the actual schedule close to the reference FIFO:

A) Receiver Participation Mechanism

1) When receiver finds that the source is transmitting out-of-order (i.e. the reference FIFO schedule is being violated), an out-of-order notification(OON) is piggy-backed by the receiver on the control packets (CTS/ACK) and it sends to the source.

2) On receiving this OON, the source goes into a back-off state after completing the transmission of its current packet.

3) The back-off period $T_{back-off}$ is given by

$$T_{back-off}=R*(EIFS+DIFS+T_{success}+CW_{min})$$

where $T_{success}$=longest possible time required to transmit a packet successfully.

B) ***Stale Entry Elimination***

1) This makes sure that the STs are free of stale entries.

2) An entry is deleted from the ST only after an ACK packet for the corresponding entry is heard by the node.

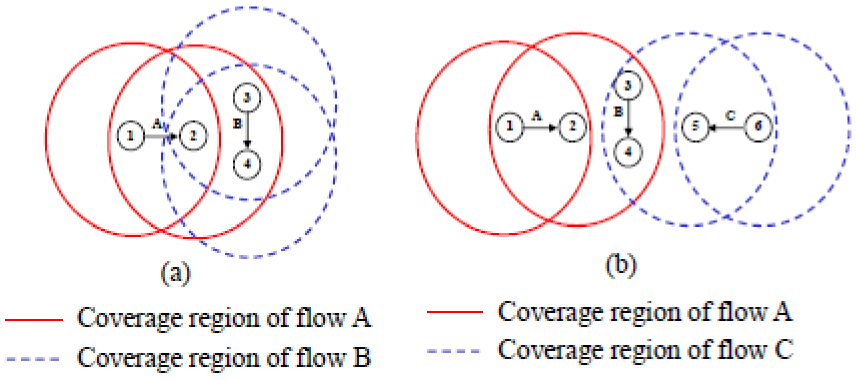

1
A
2
3
B
4
(a)
Coverage region of flow A
Coverage region of flow B
1
A
2
3
B
4
5
C
6
(b)
Coverage region of flow A
Coverage region of flow C

Figure 6.26 (a) Information asymmetry. (b) Perceived collisions.

DISTRIBUTED LAXITY BASED PRIORITY SCHEDULING SCHEME (DLPS)

• It is a packet scheduling scheme, where scheduling decisions are made taking into consideration
→ the states of neighboring nodes &
→ the feedback from destination nodes regarding packet losses

• Each node maintains following 2 tables:
1) The scheduling table(ST) contains information about
→ packets to be transmitted by the node &
→ packets overheard by the node
2) The packet delivery ratio table(PDT) contains
→ the count of DATA packets transmitted &
→ the count of ACK packets received

• Incoming packets to a node are queued in the node's input-queue according to their arrival-times (Figure 6.27).

• The scheduler
-sorts packets according to their priority values and
-inserts them into the transmission queue

• The highest priority packet from this queue is selected for transmission.

• The destination node (on receiving data packets) initiates a feedback by means of which the count of DATA packets received by it is conveyed to the source through ACK packets.

• These two pieces of information (denoted by S_i) are received by the feedback information handler (FIH).

• The FIH sends the previous state information S_{i-1} to the priority function module (PFM)

• The ULB of each packet in ST is available at the node. This information is also sent to PFM, which uses the information fed to it to calculate the priority-indices of packets in the ST.

• PDR(packet delivery ratio) of the flow at any given time is computed by

$$PDR=\frac{acksRcvd}{pktsSent}$$

• Priority index of a packet is defined as

$$PI=\frac{PDR}{M}*ULB$$

where $ULB=\frac{deadline-currentTime}{remHops}$

→ ULB is the uniform laxity budget of the packet
→ M is the user-defined parameter representing the desired packet delivery ratio for the flow
→ deadline is the end-to-end deadline target of the packet
→ currentTime denotes the current time according to the node's local clock

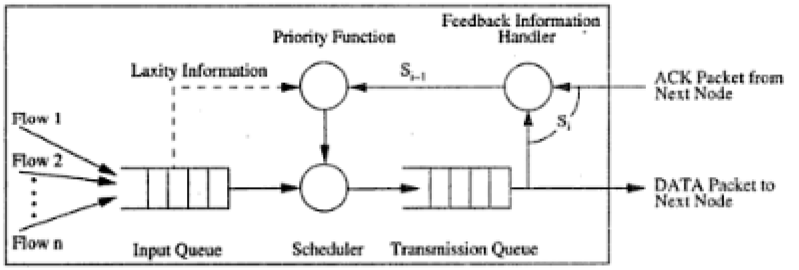


Figure 6.27. Feedback mechanism.

MAC PROTOCOLS THAT USE DIRECTIONAL ANTENNAS

MAC PROTOCOL USING DIRECTIONAL ANTENNAS

• The nodes use directional antennas for transmitting & receiving data packets, thereby reducing their interference to other neighbor nodes. This leads to an increase in the throughput of the system.

• Each node is assumed to have only one radio transceiver, which can transmit and receive only one packet at any given time (Figure 6.29).

• Each node is assumed to have 6 directional antennas.

• The protocol works as follows

1) A source transmits an RTS addressed to the receiver on all its antennas (omnidirectional transmission).
2) The intended receiver responds by transmitting CTS, again on all its antennas (omnidirectional transmission).
3) The receiver also notes down the direction of the source by identifying the antenna that receives the RTS with maximum power. The source determines the direction of the receiver in a similar manner.
4) After receiving the CTS, the source transmits the next DATA through the chosen directional antenna.
5) All other antennas are switched off and remain idle.
6) The neighbors that receive the RTS or CTS defer their transmissions for appropriate periods of time.

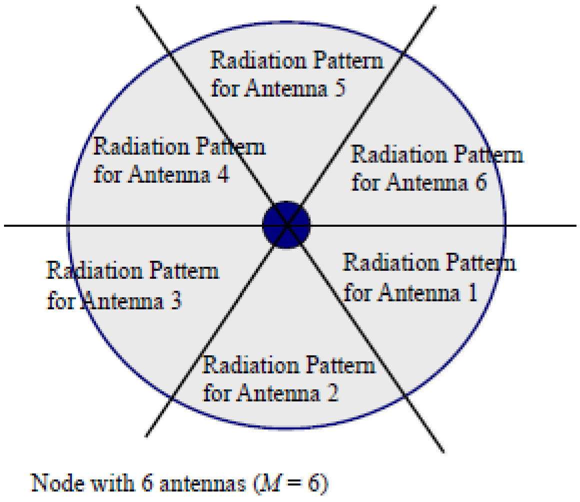

Radiation Pattern for Antenna 5
Radiation Pattern for Antenna 4
Radiation Pattern for Antenna 6
Radiation Pattern for Antenna 3
Radiation Pattern for Antenna 1
Radiation Pattern for Antenna 2
Node with 6 antennas (M = 6)

Figure 6.28. Radiation patterns of directional antennas.

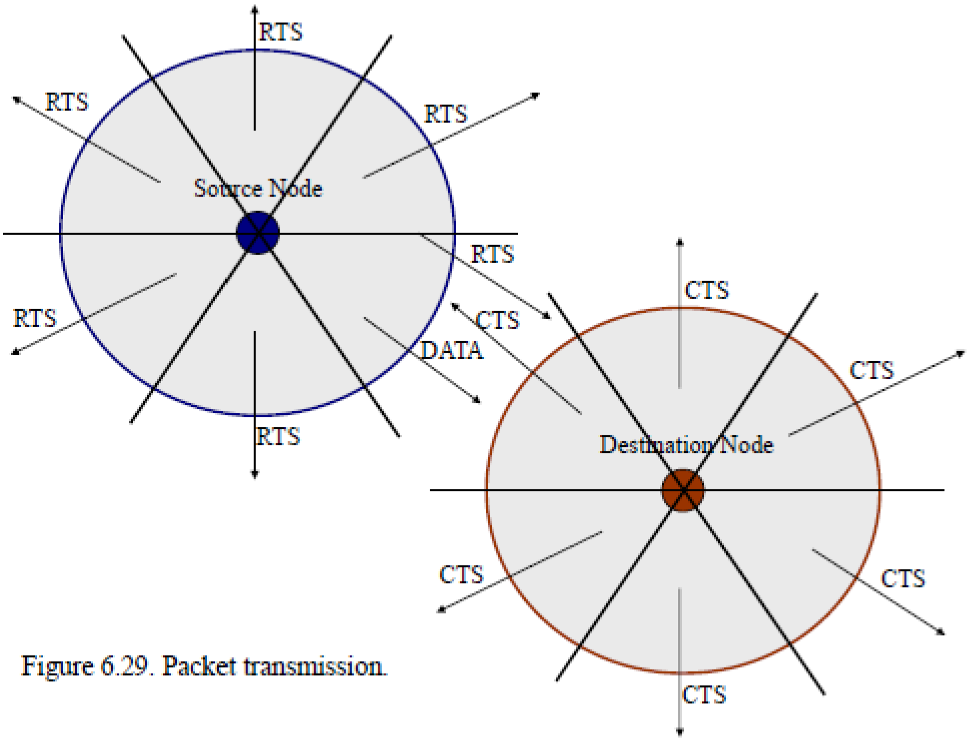


Figure 6.29. Packet transmission.

DIRECTIONAL BUSY TONE BASED MAC(BTMA) PROTOCOL

- The nodes use directional antennas for transmitting & receiving data packets, thereby reducing their interference to other neighbor nodes. This leads to an increase in the throughput of the system.
- The purpose of the busy tones (BTs) is as follows:

  1) Before transmitting an RTS, the source makes sure that the BT$_r$ tone is not active in its neighborhood, so that its transmissions do not interfere with packets being received at a neighboring receiver. Similarly, a receiver, before transmitting CTS, verifies that a BT$_t$ is not active in its neighborhood.

  2) The directional busy tones can permit simultaneous transmissions in the neighborhood of a source or a receiver (Figure 6.30).
- The protocol works as follows

  1) A source transmits an RTS addressed to the receiver on all its antennas (omnidirectional transmission).

  2) On receiving this RTS, the receiver determines the antenna-element on which the RTS is received with maximum gain.

  3) The receiver then sends back a directional-CTS(DCTS) to the source using the selected antenna-element. It also turns on busy tone BT$_r$ in the direction towards source.

  4) On receiving the CTS, the source turns on busy tone BT$_t$ in the direction towards receiver.

  5) Once the packet transmission is over, the source turns off the BT$_t$ signal.

  6) After receiving the DATA packet, the receiver turns off the BT$_r$ signal.

(For a unicast transmission, only a single antenna element is used. For broadcast transmission, all the N antenna elements transmit simultaneously.)

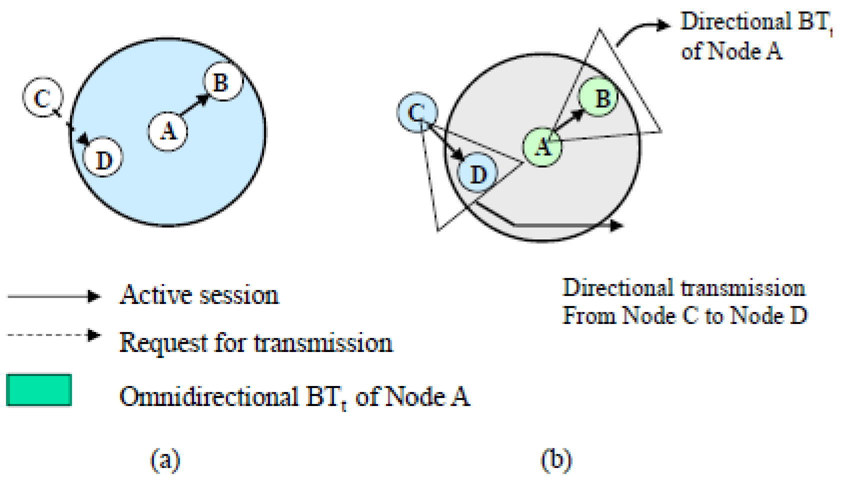

C
B
A
D
C
B
A
D
Directional BT$_t$ of Node A
Directional transmission From Node C to Node D
Active session
Request for transmission
Omnidirectional BT$_t$ of Node A
(a)
(b)

Figure 6.30. Directional DBTMA:

DIRECTIONAL MAC PROTOCOLS FOR ADHOC WIRELESS NETWORKS

- Key concept: Though a particular antenna of a node may remain blocked, the remaining antennas of the node can be used for transmissions (Figure 6.32).
- It is assumed that each node knows about the location of its neighbors as well as its own location.

DMAC-1

- A directional antenna is used for transmitting RTS, DATA & ACK.

  While an omnidirectional directional antenna is used for transmitting CTS.
- Consider figure 6.32. Here node A first transmits a directional-RTS(DRTS) to node B.
- Node B responds by transmitting an omnidirectional-CTS(OCTS).
- Then, node A sends a DATA using a directional antenna.
- When node B receives the DATA, it immediately transmits a directional-ACK(DACK).
- When node C receives OCTS from node B, only the directional antenna pointing toward node B would be blocked
- Node C can freely transmit to node D using another directional antenna.
- Drawback: The usage of DRTS may increase the probability of control packet collisions.

DMAC-2

- In DMAC-2, both directional-RTS (DRTS) as well as omnidirectional-RTS (ORTS) transmissions are used.
- A node that wants to initiate a data transfer may send an ORTS or a DRTS as per the following two rules:

  1) If none of the directional antennas at the node are blocked, then the node sends an ORTS.

  2) Otherwise, the node sends a DRTS, provided the desired directional antenna is not blocked.
- Consider figure 6.32. Here when node A initiates data transfer to node B, assuming all its antennas are not blocked, it sends an ORTS to node B.
- Node E would now receive this packet, and the antenna on which the ORTS was received would remain blocked for the duration of the transmission from node A to node B.
- If node E wants to send a packet to node A, it needs to wait for the duration of the transmission between nodes A and B.

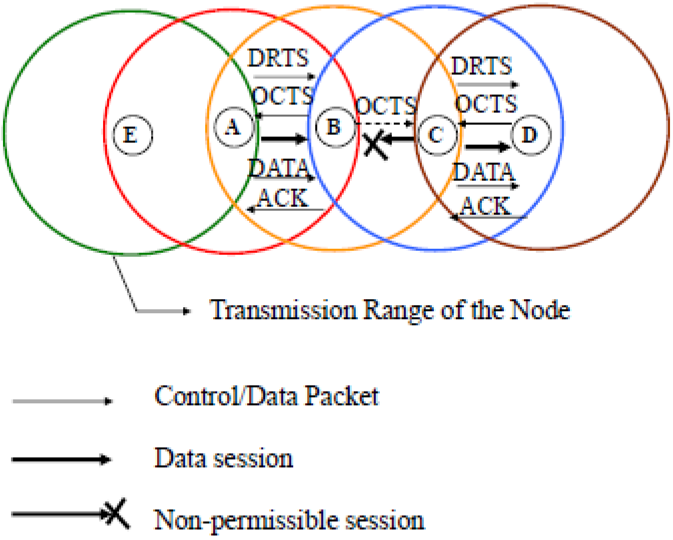


Figure 6.32. Operation of DMAC protocol.

OTHER PROTOCOLS

MULTICHANNEL MAC PROTOCOL (MMAC)

- Each node maintains a data structure called PCL (PreferableChannelList).
- PCL contains the usage of the channels within the transmission-range of the node.
- Based on their usage, channels can be classified into three types:
  1) High preference channel (HIGH): The channel has been selected by the current node and is being used by the node in the current beacon-interval.
  2) Medium preference channel (MID): The channel is free and is not being currently used in the transmission-range of the node.
  3) Low preference channel (LOW): The channel is already being used in the transmission-range of the node by other neighboring nodes. A counter is associated with each LOW state channel.
- Time is divided into beacon-intervals & every node is synchronized by periodic beacon transmissions (Fig 6.34).
- At the start of every beacon-interval, there exists a time interval called the adhoc traffic indication messages (ATIM) window.
- ATIM window is used by the nodes to negotiate for channels for transmission during the current beacon-interval.
- The protocol works as follows
  1) A source sends an ATIM to the intended receiver. The ATIM carries the PCL of the source.
  2) On receiving this ATIM, the receiver uses the PCL carried on the ATIM and its own PCL to select a channel. It includes this channel information in the ATIM-ACK packet & sends to the source.
  3) Then, source determines whether it can transmit on the channel indicated in the ATIM-ACK message.
  If so, it responds by sending the receiver an ATIM-RES(reservation) packet.
  4) At the end of the ATIM window, the source and receiver switch to the agreed-upon channel and start communicating by exchanging RTS/CTS.
- If a receiver node R receives an ATIM packet from a source S, it selects a channel as below
  1) If there exists a HIGH state channel in the node R's PCL, then that channel is selected.
  2) if there exists a HIGH state channel in the PCL of node S, then that channel is selected.
  3) if there exists a common MID state channel in the PCLs of both node S and node R, then that channel is selected.
  4) if there exists a channel which is in the MID state at only one of the two nodes, then that channel is chosen.
  5) If all channels in both PCLs are in the LOW state, the counters of the corresponding channels at nodes S and R are added, and the channel with the least count is selected.

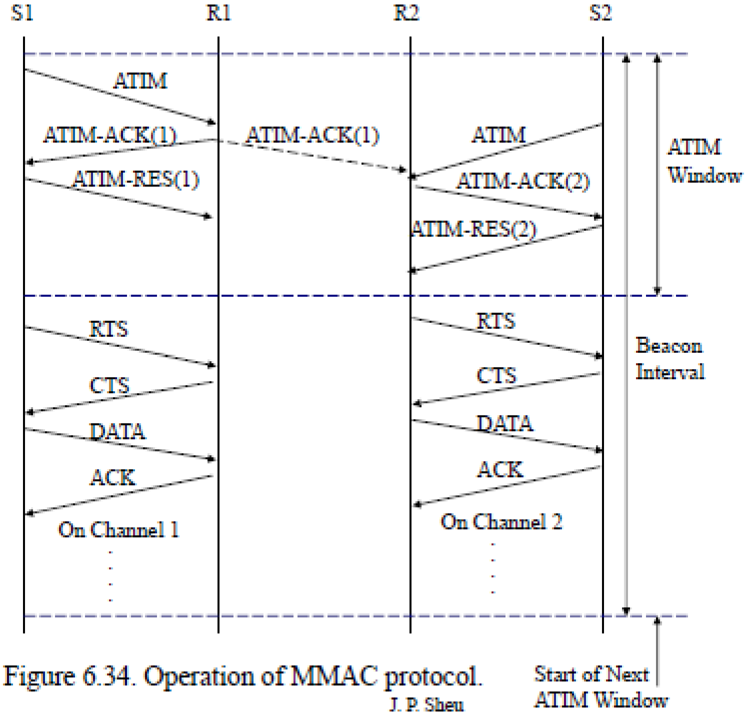


Figure 6.34. Operation of MMAC protocol.
J. P. Sheu

MULTICHANNEL CSMA MAC PROTOCOL (M-CSMA)

- It employs the notion of soft channel reservation, where preference is given to the channel that was used for the previous successful transmission.
- The available bandwidth is divided into several channels.
- The channels may be created in the frequency domain(FDMA) or in the code domain(CDMA).
- An idle node continuously monitors all the channels.
- A channel whose TRSS(total received signal strength) is below the ST(sensing threshold) of the node is marked IDLE by the node. Such channels are put in the free-channels list (FCL).
- When an idle node receives a packet to be transmitted, it does the following.

  1) If FCL is empty, it waits for any channel to become IDLE.

  2) In case FCL is non-empty, the node first checks whether the channel it used for its most recent successful transmission is included in the list. If so, the node uses this channel for its new transmission.

  3) Otherwise, one among the IDLE channels available in the FCL is randomly chosen.
- Drawback: If the number of channels is very large, then the protocol results in very high packet transmission time.

POWER CONTROL MAC PROTOCOL (PCM)

- It allows nodes to vary their transmission power levels on a per-packet basis.
- It is based on power control protocol which is referred to as the BASIC protocol.

BASIC

- The RTS and CTS packets are transmitted with maximum power p_{max} (Figure 6.35).
- The RTS-CTS handshake is used for deciding upon the transmission power for the subsequent DATA and ACK packet transmissions.
- This can be done using two methods.

First Method

1) Source A transmits the RTS with maximum power p_{max}. This RTS is received at the receiver with signal level p_r.
2) The receiver B can calculate the minimum required transmission power level $p_{desired}$ for the DATA packet based on p_r , p_{max} and the noise level at receiver B.
3) Receiver B then specifies this $p_{desired}$ in the CTS packet & sends to Source A.
4) Source A transmits the DATA packet using power level $p_{desired}$.

Second Method

1) When the receiver B receives an RTS packet, it responds with a CTS packet at the usual maximum power level p_{max}.
2) When the source A receives this CTS packet, it calculates $p_{desired}$ based on p_r & p_{max} as

$$p_{desired} = \frac{p_{max}}{p_r} * Rx_{thresh} * c$$

where Rx_{thresh}=minimum necessary received signal strength and
c=constant

3) The source uses power level $p_{desired}$ to transmit the DATA packet.

PCM

- The BASIC scheme uses

→ maximum transmit power for RTS and CTS packets, and
→ only necessary power levels for the DATA & ACK packets

- But this scheme has a drawback. Consider figure 6.35. Since the DATA & ACK transmissions use only the minimum necessary power, the DATA transmitted by node A cannot be sensed by node X. So, the packet transmitted by node X would collide at node A with the ACK packet from node B.
- PCM modifies this scheme so as to minimize the probability of collisions.
- The source and receiver nodes transmit the RTS and CTS packets with maximum power p_{max}.
- Nodes in the carrier-sensing zones of the source and receiver nodes set their NAVs for EIFS duration when they sense the signal but are not able to decode it.
- The source node generally transmits with minimum necessary power, as in the BASIC scheme.
- But, in order to avoid collisions with packets transmitted by the nodes in its carrier-sensing zone, the source node transmits the DATA packet at maximum power level p_{max} periodically (Figure 6.36).
- Since the nodes in the carrier-sensing zone defer their transmissions for EIFS duration if they are not able to decode the received signal, the transmit power for the DATA packet is increased every EIFS duration.

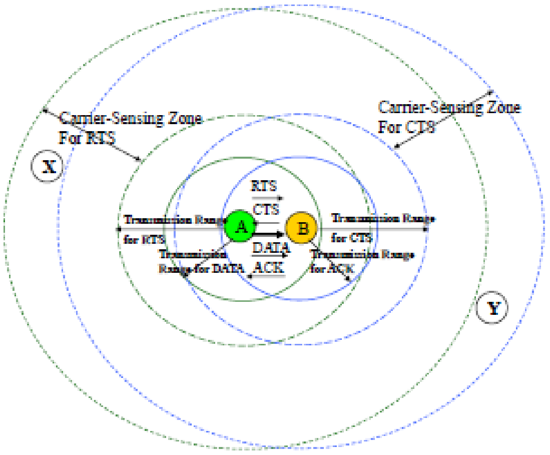


Figure 6.35. Packet transmission in BASIC scheme.

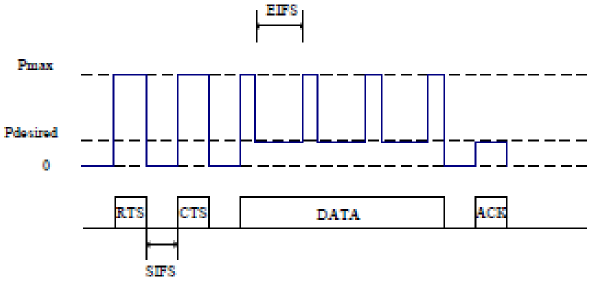


Figure 6.36 . Transmission power pattern in PCM.

RECEIVER-BASED AUTORATE PROTOCOL (RBAR)

- It uses a novel rate adaptation approach.
- *Rate adaptation* is the process of dynamically switching data rates in order to match the channel conditions so that optimum throughput for the given channel conditions is achieved.
- Rate adaptation consists of two processes, namely, channel quality estimation and rate selection.
- Rate selection is done at the receiver on a per-packet basis during the RTS-CTS packet exchange.
- The RTS and CTS packets consists of

  → the chosen data rate and

  → the size of the data packet
- The protocol works as follows

  1) The source chooses a data-rate based on some heuristic and inserts the chosen data-rate and data-size into the RTS(Figure 6.37).

  2) When a neighbor receives this RTS, it calculates the duration of reservation D_{RTS} using the data-rate and data-size carried on the RTS.

  3) While receiving the data-packet, the receiver generates an estimate of the channel conditions for the impending data transfer. Based on this estimate, it chooses an appropriate data-rate. The receiver stores the chosen data-rate and the data-size in the CTS and transmits the CTS to the source.

  4) On receiving the CTS, the source responds by transmitting the data-packet at the rate chosen by the receiver.
- Problem: If the data-rates chosen by the sender and receiver are different, then the reservation duration D_{RTS} calculated by the neighbors of the sender would not be valid. (D_{RTS} time period, which is calculated based on the information carried initially by the RTS packet, is referred to as tentative reservation)
- In order to overcome this problem, the source sends the data-packet with a special MAC header containing a RSH (reservation subheader).
- The fields in the RSH contain control information for determining the duration of the transmission.
- A neighbor node with tentative reservation entries in its NAV, on hearing the data packet, calculates D_{RSH}, the new reservation period, and updates its NAV to account for the difference between D_{RTS} & D_{RSH}.

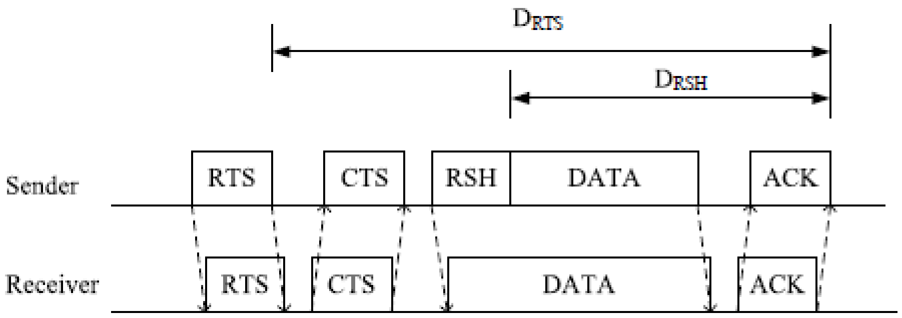


Figure 6.37. Packet transmission in RBAR

INTERLEAVED CARRIER SENSE MULTIPLE ACCESS PROTOCOL (I-CSMA)

- It efficiently overcomes the exposed terminal problem.
- Consider figure 6.38. Here, when a transmission is going on from node A to node B, nodes C and F would not be permitted to transmit to nodes D and E respectively.
- Node C is called a sender-exposed node, and node E is called a receiver-exposed node.
- The total available bandwidth is split into two equal channels (say, channel 1 and channel 2).
- The handshaking process is interleaved between the two channels, hence the name interleaved carrier-sense multiple access.
- The protocol works as follows

  1) The source transmits the RTS on channel 1(Figure 6.39).

  2) On receiving RTS, the receiver checks its E-NAV and finds out whether free time slots are available. It sends the CTS only if free slots are available. (Each node maintains a data structure called extended network allocation vector.)

  3) On receiving this CTS, the source transmits the DATA on channel 1.

  4)The receiver responds with the ACK on channel 2.
- The performance improvement is attributed to the following facts

  1) Nodes that hear RTS in a particular channel(say channel 1) and do not hear the corresponding CTS on the other channel(channel 2) concludes that they are only sender-exposed in channel 1.Therefore,if they have packets to send, they can use channel 1 to transmit RTS to other nodes.

  2) Nodes that hear only the CTS in a particular channel (say channel 1) and had not heard the corresponding RTS on the other complementary channel (channel 2) realize that they are only receiver-exposed on channel 1 to the on-going transmission. If they receive any RTS on channel 2, they would not refrain from sending CTS on channel 1 for the received RTS.

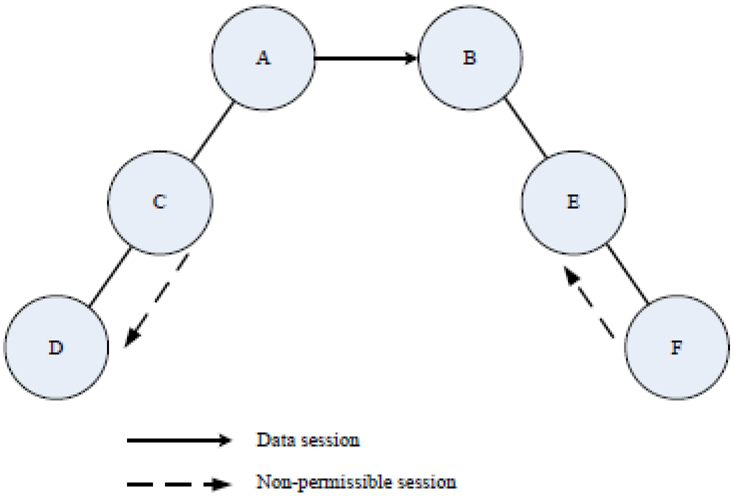


Figure 6.38. Exposed terminal problem.

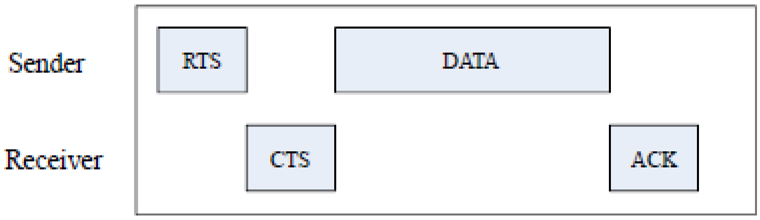


Figure 6.39. (a) Packet transmissions in 802.11 DCF

UNIT 6: TRANSPORT LAYER

ISSUES IN DESIGNING A TRANSPORT LAYER PROTOCOL FOR AD HOC WIRELESS NETWORKS

1. Induced Traffic

• In a path having multiple link, the traffic at any given link (or path) due to the traffic through neighbouring links (or paths) is referred to as *induced traffic*.

• This is due to

-the broadcast nature of the channel &

-the location-dependent contention on the channel

• This affects the throughput achieved by the protocol.

2. Induced Throughput Unfairness

• This refers to the throughput unfairness at the transport layer due to the throughput (or delay) unfairness existing at the lower layer such as the network and MAC layers.

• A transport layer should consider these in order to provide a fair share of throughput across contending flows

3. Separation of Congestion Control, Reliability and Flow Control

• The protocol can provide better performance if reliability, flow-control and congestion-control are handled separately.

• Reliability and flow-control are end-to-end activities,

whereas congestion-control can at times be a local activity.

• Objective: minimization of the additional control overhead generated by them.

4. Power & Bandwidth Constraints

• Nodes face resource constraints including the two most important resources:

(i) power source & (ii) bandwidth

• The performance of a protocol is significantly affected by these resource constraints.

5. Interpretation of Congestion

• Interpretation of network congestion as used in traditional networks is not appropriate in adhoc networks.

• This is because following parameters can also lead to packet loss:

→ high error rates of wireless channel

→ location-dependent contention

→ hidden terminal problem

→ packet collisions in the network

→ path breaks due to mobility of nodes and

→ node failure due to drained battery

6. Completely Decoupled Transport Layer

• Another challenge faced by Transport layer protocol is the interaction with the lower layers.

• Cross-layer interaction between the transport layer and lower layers is important to adapt to the changing network environment.

7. Dynamic Topology

• Experience rapidly changing network topology due to mobility of nodes.

• This leads to

→frequent path breaks

→partitioning and remerging of networks &

→high delay in re-establishment of paths

• Performance is affected by rapid changes in network topology.

DESIGN GOALS OF A TRANSPORT LAYER PROTOCOL FOR AD HOC WIRELESS NETWORKS

• The protocol should maximize the throughput per connection.

• It should provide throughput fairness across contending flows.

• It should incur minimum connection set up and connection maintenance overheads.

• It should have mechanisms for congestion control and flow control in the network.

• It should be able to provide both reliable and unreliable connections as per the requirements of the application layer.

• It should be able to adapt to the dynamics of the network such as rapid changes in topology.

• Bandwidth must be used efficiently.

• It should be aware of resource constraints such as battery power and buffer sizes and make efficient use of them.

• It should make use of information from the lower layers for improving network throughput.

• It should have a well-defined cross-layer interaction framework.

• It should maintain End-to-End Semantics.

CLASSIFICATION OF TRANSPORT LAYER SOLUTIONS

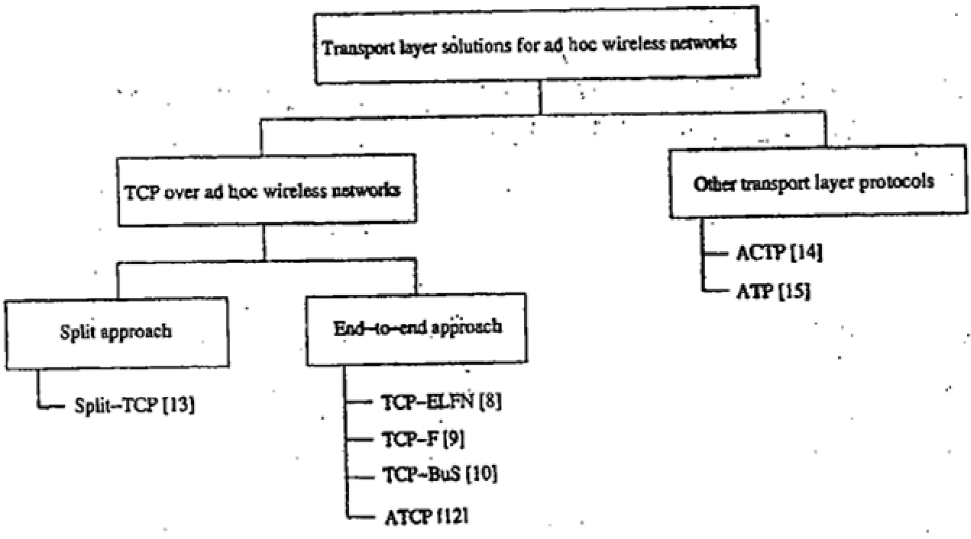

Transport layer solutions for ad hoc wireless networks
TCP over ad hoc wireless networks
Other transport layer protocols
ACTP [14]
ATP [15]
Split approach
End-to-end approach
Split-TCP [13]
TCP-ELFN [8]
TCP-F [9]
TCP-BuS [10]
ATCP [12]

Figure 9.1. Classification of transport layer solutions.

TCP OVER AD HOC WIRELESS NETWORKS

- TCP is reliable, end-to-end, connection-oriented TL protocol that provides a byte stream based service.
- Major responsibilities of TCP include
  - → Congestion control
  - → Flow control
  - → In-order delivery of packets
  - → Reliable transportation of packets

WHY TCP DOES NOT PERFORM WELL IN ADHOC WIRELESS NETWORK

1. Misinterpretation of Packet Loss

- In traditional TCP design, the packet loss is mainly attributed to network congestion.
- Adhoc network experience a much higher packets loss due to
  → High bit rate
  → Increased Collections etc.

2. Frequent Path Breaks

- If the route re-establishment time is greater than the RTO period of sender, then the sender
  → assumes congestion in the network
  → retransmits lost packets and
  → initiates congestion control algorithm
- This leads to wastage of bandwidth and battery power

3. Effect of Path Length: As path length increases, the throughput decreases. (Figure: 9.3 & 9.4)

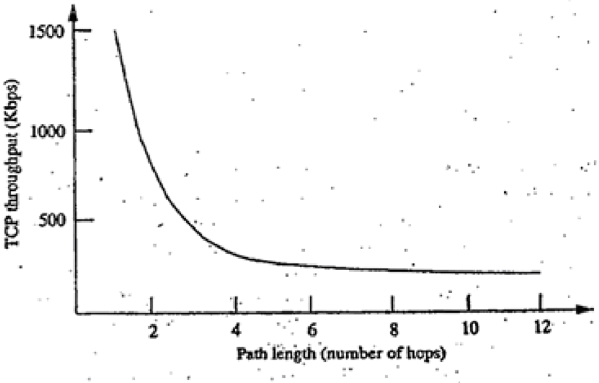


Figure 9.3. Variation of TCP throughput with path length.

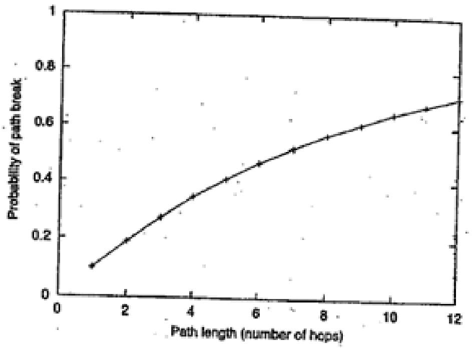


Figure 9.4. Variation of p_b with path length ($p_l = 0.1$).

4. Misinterpretation of Congestion Window

- When there are frequent path breaks, the congestion window may not reflect the maximum transmission-rate acceptable to the network and the receiver.

5. Asymmetric Link Behavior

- Radio channel has different properties such as location dependent contention, directional properties etc leading to asymmetric links.
- This can lead to TCP invoking the congestion control algorithm and several retransmissions.

6. Uni-directional Path

- TCP relies on end-to-end ACK for ensuring reliability.
- Path break on an entirely different reverse path can affect the performance of the network as much as a path breaks in the forward path.

7. Multipath Routing

- For TCP, multipath routing leads to significant amount of out-of-order packets, when intern generates a set of duplicate acknowledgement(DUPACKs), which cause additional power consumption and invocation of congestion control.

8. Network Partitioning & Remerging

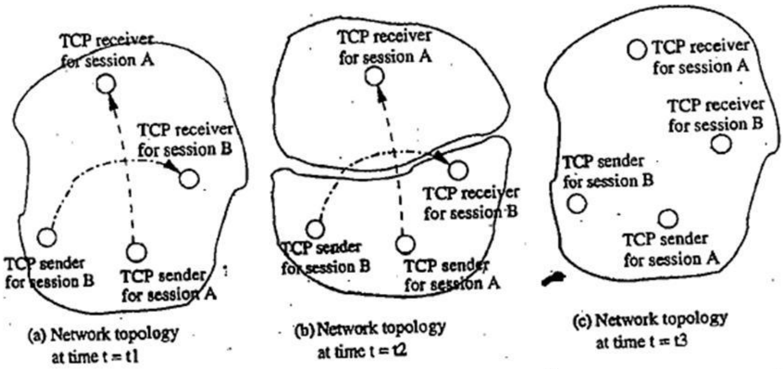

TCP receiver for session A
TCP receiver for session B
TCP sender for session B
TCP sender for session A
(a) Network topology at time t = t1
TCP receiver for session A
TCP receiver for session B
TCP sender for session B
TCP sender for session A
(b) Network topology at time t = t2
TCP receiver for session A
TCP receiver for session B
TCP sender for session B
TCP sender for session A
(c) Network topology at time t = t3

Figure 9.5. Effect of partitioning and merging of network.

- Fig:9.5 illustrate the effect of network partitions in adhoc networks.
- A network with two TCP sessions A & B is shown in (a) at time t1.
- At time t2, the network gets partitioned into two as shown in (b) due to dynamic topological changes.
- Now TCP session A's sender & receiver belong to two different partitions & TCP session B experiences path break.

FEEDBACK BASED TCP (TCP – F)

- Improves performance of TCP.
- Uses a feedback based approach.
- The routing protocol is expected to repair the broken path within a reasonable time period.

Operation

- An intermediate node, upon detection of a path break, originates route-failure-notification (RFN) packet. This intermediate-node is called Failure-point (FP).
- The RFN packet is routed toward the sender of the TCP session (Figure 9.6).
- Sender information is obtained from packets.
- If any intermediate nodes that receive RFN has an alternate route to the same destination, then it

  →discards the RFN packet &

  →uses the alternate path for forwarding further data packets, thus reducing control overhead involved in the route reconfiguration process
- When sender receives an RFN packet, it goes into a state called snooze.
- In snooze state, a sender,

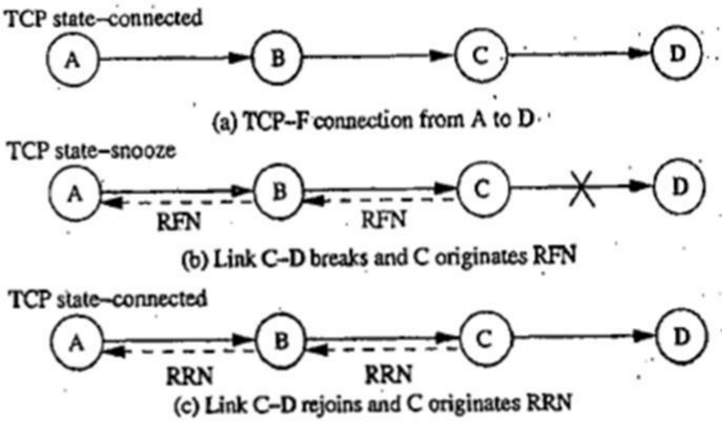
  → stops sending any more packets to the destination

  → cancels all timers

  → freezes its congestion window

  → freezes the retransmission timer

  → sets up a route failure timer
- When route failure timer expires, the sender changes from snooze-state to connected-state.
- When the route re-establishment has been done, then the failure-point sends Route Re-establishment Notification (RRN) packet to the sender and the TCP state is updated back to the connected-state.



(a) TCP-F connection from A to D

(b) Link C–D breaks and C originates RFN

(c) Link C–D rejoins and C originates RRN

Figure 9.6. Operation of TCP-F.

Advantages

- Simple feedback solution for problem arising from path breaks.
- Permits TCP congestion control mechanism to respond to congestion in the network.

Disadvantages

- If a route to sender is not available at the FP, then additional control packets may need to be generated for routing RFN packets.
- TCP-F has an additional state compared to traditional TCP state mechanism.
- Congestion window used after a new route is obtained may not reflect the achievable transmission-rate acceptable to the network and the TCP-F receiver.

TCP WITH EXPLICIT LINK FAILURE NOTIFICATION (TCP-ELFN)

- Improves TCP performance in adhoc network.
- Similar to TCP-F.

Operation

- An intermediate node, upon detection of a path break, originates ELFN packet.
- This can be implemented in two ways:
  - → by sending an ICMP Destination Unreachable (DUR) message to the sender or
  - → by piggy-backing this information to the sender
- Once the sender receives the ELFN packet, it disables its retransmission timers and enters a standby state.
- In standby state, it periodically originates probe packets to see if a new route is established.
- Upon reception of an ACK by the receiver for the probe packets, it leaves the standby state, and continues to function as normal.

Advantages

- Improves TCP performance by decoupling the path break information from the congestion information by the use of ELFN.
- Less dependent on routing protocol & requires only link failure notification about the path break.

Disadvantages

- When the network is temporarily partitioned, the path failure may last longer & this can lead to the origination of periodic probe packets consuming bandwidth & power.
- Congestion window used after a new route is obtained may not reflect the achievable transmission-rate acceptable to the network and the TCP receiver.

TCP-BUS (TCP WITH BUFFERING CAPABILITY AND SEQUENCE INFORMATION)

- It is similar to TCP-F and TCP-ELFN in its use of feedback information from an intermediate-node on detection of a path break. But it is more dependent on the routing protocol.
- TCP-BuS was proposed, with Associativity-Based Routing (ABR) protocol as the routing scheme.

Operation

- An upstream intermediate node, upon detection of a path break, originates ERDN packet to the TCP-sender. This intermediate-node is called pivot-node (PN).
- Upon receiving ERDN packet, the TCP-sender
  →stops transmission &
  →freezes all timers and windows as in TCP-F
- The packets in transmit at the intermediate nodes from the TCP-sender to the PN are buffered until a new partial path from the PN to the TCP-receiver is obtained by the PN.
- The downstream node, upon detection of a path break, originates Route Notification (RN) packet to TCP-receiver.
- PN attempts to find new partial path (route) to the TCP-receiver, and the availability of such a partial path to destination is intimated to the sender through an explicit route successful notification (ERSN) packet.
- TCP utilizes route reconfiguration mechanism of ABR to obtain partial path to the destination.
- TCP-sender also periodically originates probe packets to check the availability of a path to the destination.

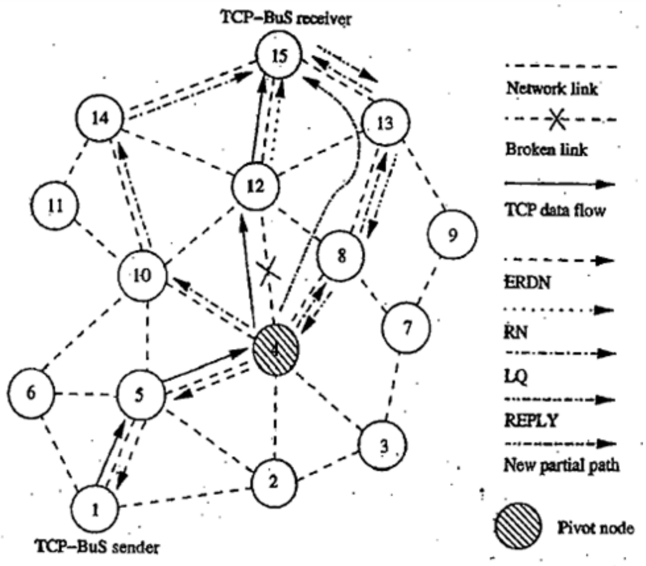


Figure 9.7. Operation of TCP-BuS.

Advantages

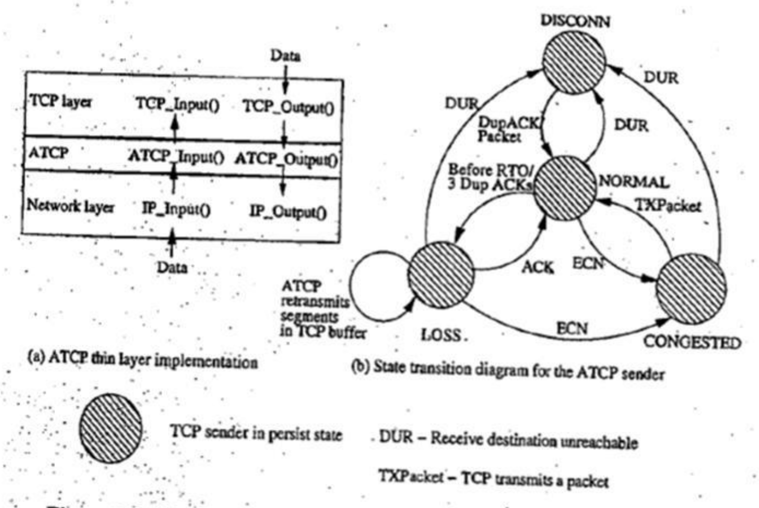
- Performance improvement.
- Avoidance of fast retransmission due to the use of buffering, sequence numbering, and selective acknowledgement
- Also takes advantage of the underlying routing protocols.

Disadvantages

- Increased dependency on the routing protocol and the buffering at the intermediate nodes.
- The failure of intermediate nodes that buffer the packets may lead to loss of packets and performance degradation.
- The dependency on the routing protocol may degrade its performance with order routing protocols that do not have similar control messages as in ABR.

ADHOC TCP

- Based on feedback information received from the intermediate nodes, the TCP sender changes its state to the
  - → Persist state
  - → Congestion control state or
  - → Retransmission state
- When an intermediate node finds that the network is partitioned, then the TCP sender state is changed to the persist state where it avoids unnecessary retransmissions.
- Figure shows the thin layer implementation of ATCP between the traditional TCP layer and the IP layer.
- This does not require changes in the existing TCP protocol.
- This layer is active only at the TCP sender.



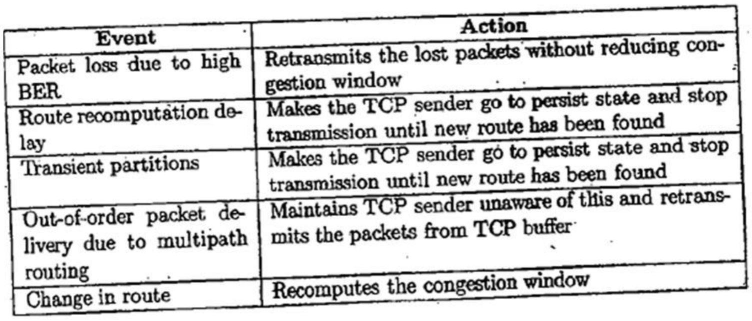
(a) ATCP thin layer implementation

(b) State transition diagram for the ATCP sender

Figure 9.8: An illustration of ATCP thin layer and ATCP state diagram.

- Major function of the ATCP Layer is that it monitors the:
  - → Packet sent & received by TCP sender
  - →The state of the TCP sender
  - → State of the network
- The four states in the ATCP are:
  1. NORMAL.
  2. CONGESTED
  3. LOSS
  4. DISCONN
- When a TCP connection is established, the ATCP sender state is in NORMAL, here ATCP does not interfere with the operation of TCP and it remains invisible.

Table 9.1. The actions taken by ATCP



Event	Action
Packet loss due to high BER	Retransmits the lost packets without reducing congestion window
Route recomputation delay	Makes the TCP sender go to persist state and stop transmission until new route has been found
Transient partitions	Makes the TCP sender go to persist state and stop transmission until new route has been found
Out-of-order packet delivery due to multipath routing	Maintains TCP sender unaware of this and retransmits the packets from TCP buffer
Change in route	Recomputes the congestion window

Advantages

- It maintains the end to end semantics of TCP.
- It is compatible with traditional TCP.
- Improves throughput of TCP in adhoc wireless network.

Disadvantages

- Dependency on he network layer protocol to detect the route changes and partitions.
- Addition of thin ATCP layer to TCP/IP protocol stack requires changes in the interface functions currently being used.

SPLIT TCP

- Major issues that affect the performance of TCP over adhoc network is: the degradation of throughput with increasing path length.
- Split TCP provides a unique solution to this problem by splitting the transport layer objectives into:
  → Congestion control
  → End to End reliability
- In addition, split TCP splits a long TCP connection into a set of short concatenated TCP connections (called segments or zones) with a number of selected intermediate nodes (known as proxy nodes) as terminating points of these short connections.
- A proxy node
  → receives the TCP packets
  → reads its contents
  → stores it in its local buffer and
  → sends Local acknowledgement(LACK) to the source (or the previous proxy)
- LACK does not guarantee end to end delivery.
- The responsibility of further delivery of packets is assigned to the proxy node.
- The number of proxy nodes in a TCP session is determined by
  the length of the path between source & destination node.
- Based on a distributed algorithm, the intermediate nodes that receive TCP packets determine whether to act as a proxy node or just as a simple forwarding node.

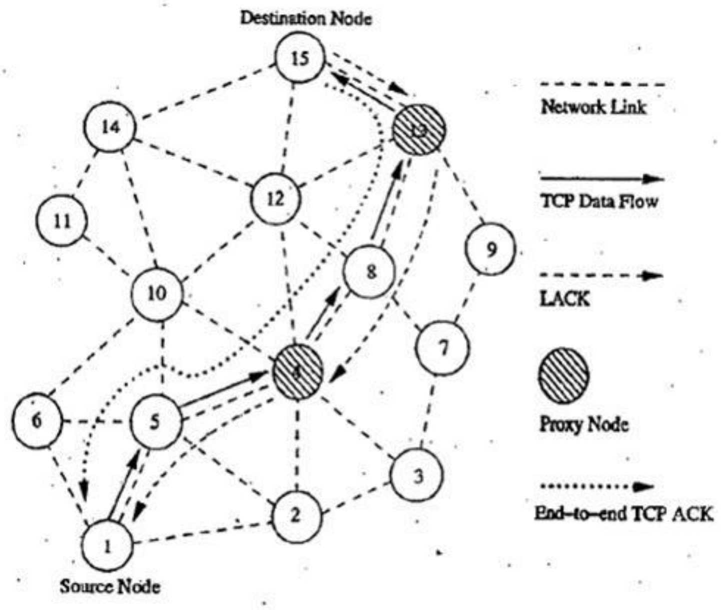


Figure 9.9. An illustration of Split-TCP.

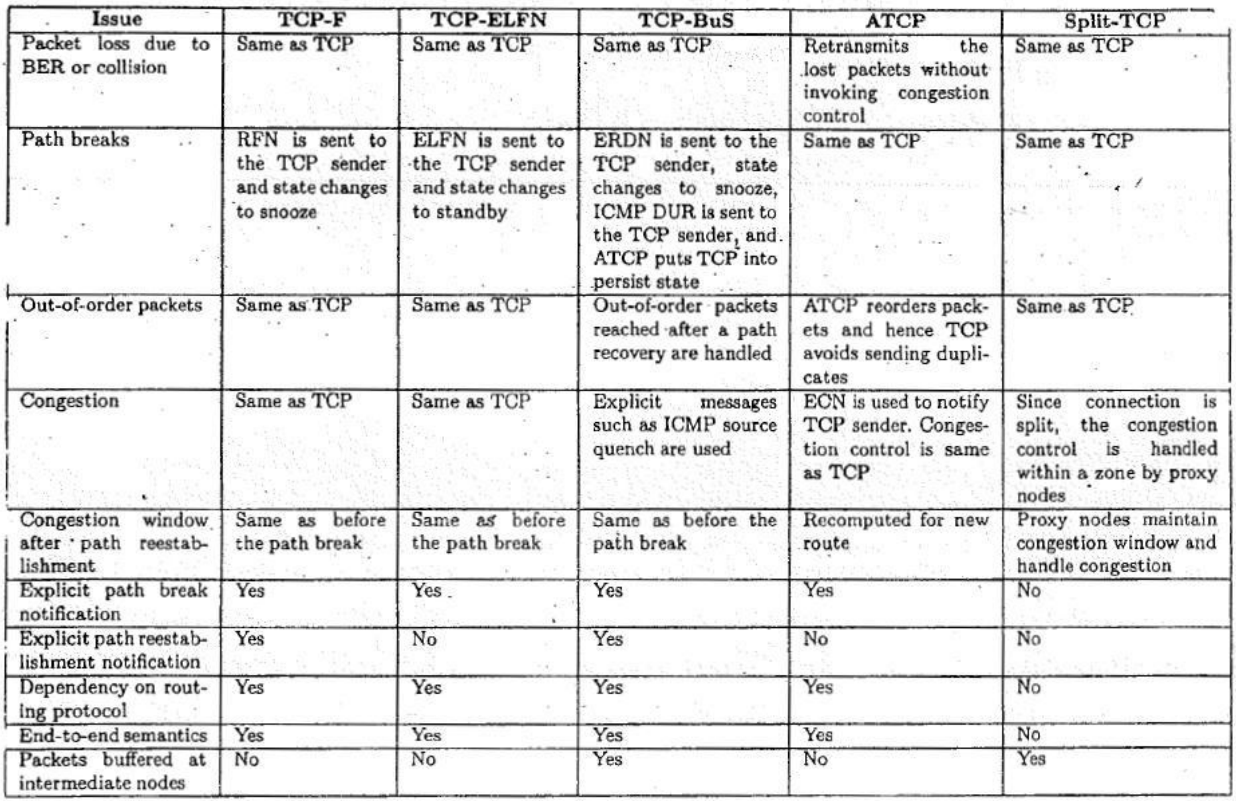
Advantages:

- Improved throughput.
- Improved throughput fairness.
- Lessened impact of mobility.

Disadvantages:

- Requires modifications to TCP protocol.
- End to End connection handling of traditional TCP is violated.
- The failure of proxy nodes can lead to throughput degradation.

COMPARISION OF TCP SOLUTIONS FOR ADHOC WIRELESS NETWORKS



Issue	TCP-F	TCP-ELFN	TCP-BuS	ATCP	Split-TCP
Packet loss due to BER or collision	Same as TCP	Same as TCP	Same as TCP	Retransmits the lost packets without invoking congestion control	Same as TCP
Path breaks	RFN is sent to the TCP sender and state changes to snooze	ELFN is sent to the TCP sender and state changes to standby	ERDN is sent to the TCP sender, state changes to snooze, ICMP DUR is sent to the TCP sender, and ATCP puts TCP into persist state	Same as TCP	Same as TCP
Out-of-order packets	Same as TCP	Same as TCP	Out-of-order packets reached after a path recovery are handled	ATCP reorders packets and hence TCP avoids sending duplicates	Same as TCP
Congestion	Same as TCP	Same as TCP	Explicit messages such as ICMP source quench are used	ECN is used to notify TCP sender. Congestion control is same as TCP	Since connection is split, the congestion control is handled within a zone by proxy nodes
Congestion window after path reestablishment	Same as before the path break	Same as before the path break	Same as before the path break	Recomputed for new route	Proxy nodes maintain congestion window and handle congestion
Explicit path break notification	Yes	Yes	Yes	Yes	No
Explicit path reestablishment notification	Yes	No	Yes	No	No
Dependency on routing protocol	Yes	Yes	Yes	Yes	No
End-to-end semantics	Yes	Yes	Yes	Yes	No
Packets buffered at intermediate nodes	No	No	Yes	No	Yes

OTHER TRANSPORT LAYER PROTOCOLS FOR AD HOC WIRELESS NETWORKS

APPLICATION CONTROLLED TRANSPORT PROTOCOL (ACTP)

- It is a light-weight transport layer protocol.
- It assigns the responsibility of ensuring reliability to the application layer.
- It stands in between TCP and UDP where
  - → TCP experiences low performance with high reliability &
  - → UDP provides better performance with high packet loss in Adhoc wireless networks.
- The key design philosophy is to
  - → leave the provisioning of reliability to the application layer &
  - → provide simple feedback information about the delivery status of packets to the application layer
- It supports the priority of packets to be delivered.
- Each API function call to send a packet contains information such as
  - → maximum delay
  - → message number &
  - → priority of the packet
- Delivery status is maintained at the ACTP layer. This reflect
  - → Successful delivery of the packet
  - → Possible loss of the packet
  - → Remaining time for the packet
  - → No state information exists at the ACTP layer

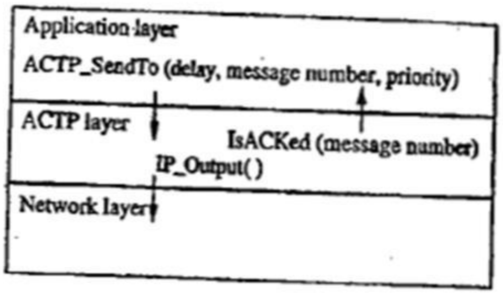


Figure 9.10. An illustration of the interface functions used in ACTP.

Advantages:

- Provides freedom of choosing the required reliability level to the application layer.
- Scalable for large networks.
- Throughput is not affected by path breaks.

Disadvantage:

- Not compatible with TCP.

ADHOC TRANSPORT PROTOCOL (ATP)

- It is specifically designed for adhoc networks and is not a variant of TCP.
- It defers from TCP in following major aspects:
  - → Co-ordination among multiple layers
  - → Rate-based transmissions
  - → Decoupling congestion control & reliability
  - → Assisted congestion control
- It uses services from network & MAC layers for improving its performance.
- It uses information from lower layers for
  - → Estimation of the initial transmission-rate
  - → Detection, avoidance and control of congestion
  - → Detection of path breaks
- It utilizes timer-based transmission.
- The network congestion information is obtained from the intermediate nodes.
- Field in which delay information is included is referred as rate feedback field.
- It has three phases namely: increase, decrease and maintain.

Advantages:

- Improved performance.
- Decoupling congestion control and reliability mechanisms.
- Avoidance of congestion window fluctuations.

Disadvantage:

- Lack of interoperability with TCP.

UNIT 7: SECURITY

NETWORK SECURITY REQUIREMENTS

A security protocol for adhoc networks should satisfy the following requirements

1. Confidentiality

- The data sent by the sender must be understandable only to the intended-receiver.
- Though an intruder might get hold of the data being sent, he must not be able to derive any useful information out of the data.
- *Data encryption* can be used to ensure confidentiality.

2. Integrity

- The data sent by the source-node should reach the destination-node without being altered.
- It should not be possible for any malicious-node to tamper with the data during transmission

3. Availability

- The network should remain operational all the time.
- It must be
  →robust enough to tolerate link-failures &
  →capable of surviving various attacks mounted on it
- It should be able to provide guaranteed services whenever an authorized-user requires them

4. Non-Repudiation

- It is a mechanism to guarantee
  →that the sender of a message cannot later deny having sent the message &
  →that the recipient cannot deny having received the message
- *Digital signatures* are used for this purpose.

ISSUES AND CHALLENGES IN SECURITY PROVISIONING

1. Shared Broadcast Radio Channel

- The radio channel used for communication
  →is broadcast in nature &
  →is shared by all nodes within its direct transmission range.
- Data transmitted by a node is received by all nodes within its direct transmission range. So, a malicious-node could easily obtain transmitted-data in the network.
- This problem can be minimized to a certain extent by using *directional antennas.*

2. Limited Resource Availability

- Resources such as bandwidth, battery-power, & computational-power are limited in adhoc networks.
- Hence, it is difficult to implement complex cryptography-based security mechanisms in networks.

3. Insecure Operational Environment

- The operating environments where adhoc wireless is used may not always be secure.
- One important application of such networks is in *battlefields.*

4. Physical Vulnerability

- Nodes in the networks are usually compact & hand-held in nature.
- They could get damaged easily & are also vulnerable to theft.

5. Lack of Central Authority

- In wired-networks & infrastructure-based wireless networks, it would be possible to
  →monitor the traffic on the network through certain important central points &
  →implement security mechanisms at such points
- Since adhoc networks do not have central points, these mechanisms cannot be applied in adhoc networks.

6. Lack of Associations

- Since these networks are dynamic in nature, a node can join or leave the network at any point of time.
- If no proper authentication mechanism is used for associating nodes in a network, an intruder would be able to
  →join into the network quite easily &
  →carry out his attacks

NETWORK SECURITY ATTACKS

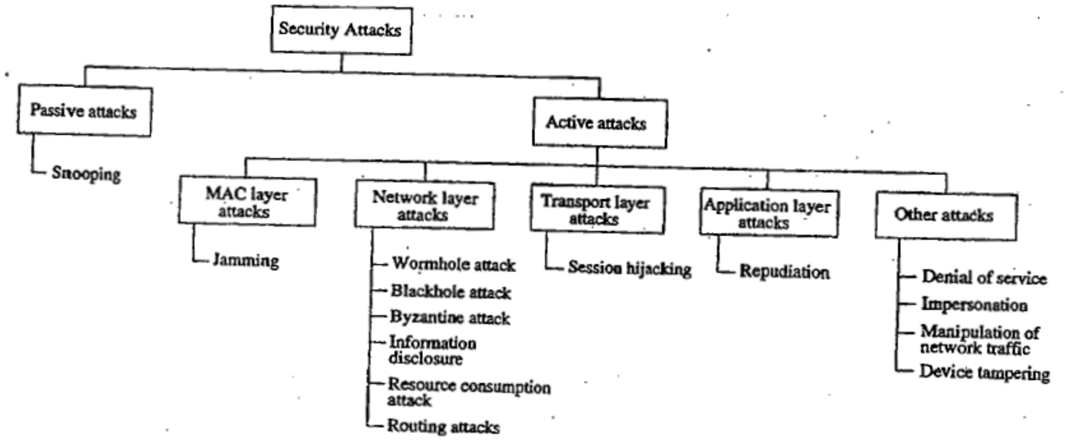


Figure 9.11. Classifications of attacks.

Attacks on adhoc networks can be classified into 2 broad categories, namely:

1. Passive Attack

- It does not disrupt the normal operation of the network (Figure: 9.1),
- The adversary snoops the data exchanged in the network without altering it.
- *Data encryption* can be used to overcome this problem.

2. Active Attack

- It disrupts the normal functioning of the network.
- It attempts to alter (or destroy) the data being exchanged in the network,
- They can be further classified into 2 categories:

  → **External attacks** are carried out by nodes that do not belong to the network. They can be prevented using standard encryption techniques and firewalls

  → **Internal attacks** are from compromised-nodes that are actually part of the network

NETWORK LAYER ATTACKS

There are many types of attacks pertaining to the network layer in network protocol stack. Some of them are as follows:

1. Wormhole Attack

• An attacker

→receives packets at one location in the network &

→tunnels them to another location in the network, where the packets are resent into the network. This tunnel between 2 colliding attackers is referred to as a *wormhole*.

• If proper mechanisms are not employed to defend the network against wormhole attacks, existing routing protocols for adhoc networks may fail to find valid routes.

2. Blackhole Attack

• A malicious-node falsely advertises good paths to destination-node during path-finding process .

• The intention of malicious-node could be

→to hinder the path-finding process or

→to intercept all data packets being sent to the destination node

3. Byzantine Attack

• A set of compromised-nodes work in collusion & carries out attack such as

→creating routing loops

→routing packets on non-optimal paths &

→selectively dropping packets

4. Information Disclosure

• A compromised-node may leak confidential information to unauthorized-nodes in the network.

5. Resource Consumption Attack

• A malicious-node tries to consume(or waste) resources of other nodes present in the network.

• The resources targeted are battery-power, bandwidth & computational-power.

6. Routing Attacks

• There are several types of attacks mounted on routing protocol & they are as follows:

i. Routing Table Overflow

• Adversary-node advertises routes to non-existent nodes, to the authorized-nodes present in the network.

• The main objective is to cause an overflow of routing-tables, which would in turn prevent the creation of entries corresponding to new routes to authorized-nodes.

ii. Routing Table Poisoning

• The compromised-nodes

→send wrong routing updates or

→modify genuine route update packets

• This may result in

→sub-optimal routing

→congestion in network or

→even make some parts of network inaccessible

iii. Packet Replication

• An adversary-node would replicate state packets.

iv. Route Cache Poisoning

• Similar to routing-table poisoning, an adversary can also poison the route cache to achieve similar activities.

v. Rushing Attack

• On-demand routing protocols that use duplicate suppression during the route discovery process are vulnerable to this attack.

TRANSPORT LAYER ATTACKS

Session Hijacking

• An adversary takes control over a session between 2 nodes.

• Since most authentication processes are carried out only at the start of session,

once the session between 2 nodes get established,

the adversary-node masquerades as one of the end-nodes of the session & hijacks the sessions.

APPLICATION LAYER ATTACKS

Repudiation

• It refers to the attempted denial by a node involved in a communication of having participated in all or part of the communication

OTHER ATTACKS

This section discusses security attacks that cannot strictly be associated with any specific layer in the network protocol stack

MULTI-LAYER ATTACKS

Multi-layer attacks are those that could occur in any layer of the network protocol stack. Some of the multi-layer attacks in adhoc networks are:

1. Denial of Service (DoS)

- An adversary attempts to prevent authorized-users of services offered by the network from accessing those services.
- This may lead to a failure in the delivery of guaranteed services to the end-users.
- Some of the DoS attacks are as follows:

  →**Jamming:** Adversary initially keeps monitoring the wireless medium in order to determine the frequency at which the receiver-node is receiving signals from the sender-node. Frequency hopping spread spectrum(FHSS) and direct sequence spread spectrum(DSSS) are two commonly used techniques that overcome jamming attacks

  →**SYN Flooding:** An adversary sends a large number of SYN packets to a victim-node, spoofing the return addresses of the SYN packets. The victim-node builds up a table for holding information regarding all pending connections. Since the maximum possible size of the table is limited, the increasing number of half-connections results in an overflow in the table.

  →**Distributed DoS Attack:** Several adversaries that are distributed throughout the network collide and prevent authorized-users from accessing the services offered by the network.

2. Impersonation

- An adversary assumes the identity & privileges of an authorized-node, either

  → to make use of network-resources that may not be available to it under normal circumstances or

  → to disrupt the normal functioning of the network
- A *man-in-the-middle* attack is another type of impersonation attack.

DEVICE TAMPERING

- Unlike nodes in a wired network, nodes in adhoc networks are usually compact, soft and hand-held in nature.
- They could get damaged or stolen easily.

KEY MANAGEMENT

• **CRYPTOGRAPHY** is one of the most common & reliable means to ensure security & can be applied to any communication network.

• The original information to be sent from one person to another is called ***plaintext***.

• The plaintext is converted into ***ciphertext*** by the process of ***encryption.***

• An authentic-receiver can decrypt the ciphertext back into plaintext by the process of ***decryption.***

• The process of encryption & decryption are governed by keys. Keys are small amounts of information used by the cryptographic-algorithms. When the keys are to be kept secret to ensure the security of the system, it is called a ***secret key.***

• The secure administration of cryptographic keys is called ***Key Management.***

• The 4 main goals of cryptography are confidentiality, integrity, authentication & non-repudiation.

• There are 2 major kinds of cryptographic algorithms:

→ ***Symmetric*** *key algorithms*, which use the same key for encryption & decryption

→ ***Asymmetric*** *key algorithms*, which use two different keys for encryption & decryption

• The asymmetric key algorithms are based on some mathematical principles which make it feasible or impossible to obtain one key from another; therefore, one of the keys can be made public while the others is kept secret (private). This is called public key cryptography.

SYMMETRIC KEY ALGORITHMS

• These rely on the presence of shared key at both the sender & receiver, which has been exchanged by some previous arrangement.

• There are 2 kinds of symmetric-key algorithms:

→ Block ciphers &

→ Stream ciphers

• A block cipher is an encryption scheme in which plaintext is broken into fixed-length segments called blocks, & the blocks are encrypted one at a time.

• The simplest example includes substitution & transposition.

• In *substitution,* each alphabet of plaintext is substituted by another in the cipher text, & this table mapping of the original & the substituted alphabet is available at both the sender & receiver.

• A *Transposition cipher* permutes the alphabet in plaintext to produce the cipher text.

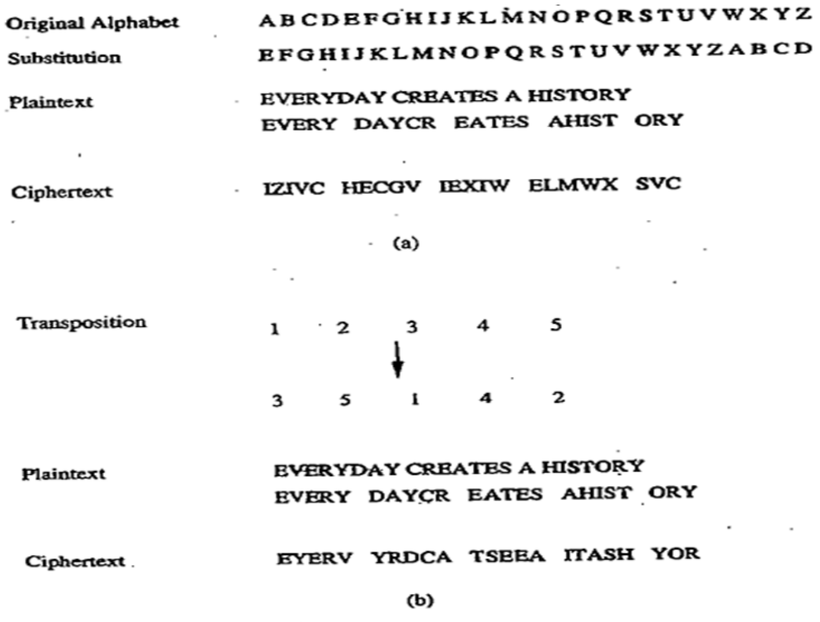

Original Alphabet
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Substitution
E F G H I J K L M N O P Q R S T U V W X Y Z A B C D
Plaintext
EVERYDAY CREATES A HISTORY
EVERY DAYCR EATES AHIST ORY
Ciphertext
IZIVC HECGV IEXIW ELMWX SVC
(a)
Transposition
1 2 3 4 5
3 5 1 4 2
Plaintext
EVERYDAY CREATES A HISTORY
EVERY DAYCR EATES AHIST ORY
Ciphertext
EYERV YRDCA TSEEA ITASH YOR
(b)

Figure 9.12. Substitution and transposition.

• Fig (a) shows encryption using substitution & fig (b) shows a transposition cipher.

• The block length used is 5.

• A stream cipher is, in effect, a block cipher of block length one.

• One of the simplest stream ciphers is *vernam cipher*, which uses a key of same length as plaintext for encryption.

• For example : If the plaintext is the binary string 10010100 & key is 01011001.then the encrypted string is given by the XOR of the plaintext & key, to be 11001101. The plaintext is again recovered by XOR-ing the cipher text with the same key.

ASYMMETRIC KEY ALGORITHMS

- These use different keys at the sender-end & receiver-ends for encryption & decryption, respectively.
- Let the encryption process be represented by a function E, & decryption by D.
  Then plaintext ‘m’ is transformed into the ciphertext ‘c’ as

$$C = E(m).$$

  The receiver then decodes c by applying D. Hence, D is such that

$$m = D(c) = D(E(m))$$

- The key E is made public, while D is made private, known only to the intended receiver.
- RSA algorithm is the best example of public key cryptography.
- **Digital signatures(DS)** scheme are also based on public key encryption.
  →In DS, the person who wishes to sign a document encrypts it using his private key D, which is known only to him.
  →Anybody who has his public key E can decrypt it and obtain the original document
  →A trusted third party is responsible for issuing these digital signatures and for resolving any disputes regarding the signatures
  →This is usually a governmental or business organization.

KEY MANAGEMENT APPROACHES

- The primary goal of key management is to share a secret among a specified set of participants.
- The main approaches to key management are key pre-distribution, key transport, key arbitration and key agreement.

1. KEY PREDISTRIBUTION

- This involves distributing key to all interested parties before the start of communication.
- This method involves much less communication & computation, but all participants must be known *a priori*, during the initial configuration.
- Once deployed, there is no mechanism to include new members in the group or to change the key.
- As an improvement over pre-distribution scheme, sub-groups may be formed within a group, and some communication may be restricted to a subgroup.

2. KEY TRANSPORT

- One of the communicating-entities generates keys & transports them to the other members.
- The simplest scheme assumes that a shared key already exists among the participating-members. This shared key is used to encrypt a new key & is transmitted to all corresponding nodes.
- Only those nodes which have the prior shared key can decrypt it.
- This is called the *Key Encrypting Key (KEK) method.*
- An interesting method for key transport without prior shared keys is the ***shamir's three-pass protocol***.(Fig: 9.13),
- The scheme is based on a special-type of encryption called communicative Encryption schemes.
- Consider 2 nodes X & Y which wish to communicate. Node X selects a key K which it wants to use in its communication with node Y. It then generates a random key K_x ,using which it encrypts K with f, & sends to node Y. Node Y encrypts this with a random key K_y using g,& sends this back to node X.
- Now, node X decrypts this message with its key K_x, & after applying inverse function f^{-1}, sends it to node y. finally, node Y decrypts the message using K_y & g^{-1} to obtain key K.

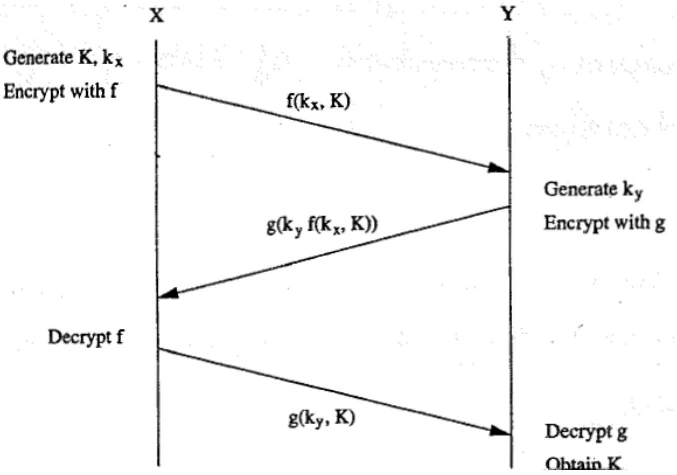


Figure 9.13: Shamir's three-pass protocol.

3. KEY ARBITRATION

- This uses a central arbitrator to create & distribute keys among all participants. Hence, they are a class of key transport schemes.
- In adhoc networks, the problem with implementation of arbitrated protocols is that the arbitrator has to be powered on at all times to be accessible to all nodes.
- This leads to a power drain on that particular node.
- Alternative is to make the keying service distributed.
- If any one of the replicated arbitrators is attacked, the security of the whole system breaks down.

4. KEY AGREEMENT

- Key agreement protocols are used to establish a secure context over which a session can be run, starting with many parties who wish to communicate & an insecure channel.
- In group key agreement schemes, each participant contributes a part to the secret key.
- Require least amount of pre-configuration
- Have high computational capability
- The most popular key agreement schemes use the Diffie-Hellman exchange, an asymmetric key algorithm based on discrete logarithms.

KEY MANAGEMENT IN ADHOC NETWORKS

- Adhoc networks pose certain specific challenges in key management, due to the lack of infrastructure in such networks.
- 3 types of infrastructure have been identified, which are absent in adhoc networks:
  - →The first is the network infrastructure, such as dedicated routers & stable links, which ensure communication with all nodes.
  - →The second missing infrastructure is services, such as name resolution, directory & TTP's.
  - →The third missing infrastructure is the administrative support of certifying authorities.

Password-Based Group Systems

- A long string is given as the password for users for one session.
- However, human beings tend to favour natural language phrases as passwords, over randomly generated strings.
- Such passwords, if used as keys directly during a session, are very week & open to attack directly during a high redundancy, & the possibility of reuse over different sessions.
- Hence, protocols have been proposed to derive a strong key (not vulnerable to attacks).
- This password-based system could be two-party, with a separate exchange between any 2 participants, or it could be for the whole group, with a leader being elected to preside over the session.
- The protocol used is as follows:
  - →Each participant generates a random number, & sends it to all others
  - →When every node has received the random number of every other node, a common pre-decided function is applied on all the numbers to calculate a reference value
  - →The nodes are ordered based on the difference between their random number & the reference value

Threshold Cryptography

- Public Key Infrastructure(PKI) enables the easy distribution of keys & is a scalable method.
- Each node has a public/private key pair.
- A certifying authority(CA) can bind the keys to a particular node.
- But CA has to be present at all times, which may not be feasible in Adhoc networks.
- A scheme based on threshold cryptography has been proposed by which n servers exist in an adhoc network, out of which any (t+1) servers can jointly perform arbitration or authorization successfully, but t servers cannot perform the same. This is called an (n, t+1) configuration, where n >= 3t +1.
- To sign a certificate, each server generates a partial signature using its private key & submits it to a combiner. The combiner can be any one of the servers.
  - →In order to ensure that the key is combined correctly, t+1 combiners can be used to account for at most t malicious servers.
  - →Using t+1 partial signatures, the combiner computes a signature & verifies its validity using a public key.
  - →If verification fails, it means that at least one of the t+1 keys is not valid, so another subset of t+1 partial signature is tried. If combiner itself is malicious, it cannot get a valid key, because partial key itself is always invalid.

Self-Organized Public Key Management for Mobile Adhoc Networks

- This makes use of absolutely no infrastructure.
- The users in the adhoc network issue certificates to each other based on personal acquaintance.
- A certificate is binding between a node & its public-key.
- The certificates are stored & distributed by the users themselves.
- Certificates are issued only for specific period of time, before it expires; the certificate is updated by the user who had issued the certificate.
- Each certificate is initially stored twice, by the issuer & by the person for whom it is issued.
- If any of the certificates are conflicting (e.g: the same public key to different users, or the same user having different pubic keys), it is possible that a malicious-node has issued a false certificate.
- A node then enables such certificates as conflicting & tries to resolve the conflict.
- If the certificates issued by some node are found to be wrong, then that node may be assumed to be malicious.
- A certificate graph is a graph whose vertices are public keys of some nodes and whose edges are public key certificates issued by users.

SECURE ROUTING IN AD HOC WIRELESS NETWORKS

• Ensuring secure communication in adhoc networks include the mobility of nodes, a promiscuous mode of operation, limited processing power & limited availability of resources such as battery power, bandwidth & memory.

REQUIREMENTS OF A SECURE ROUTING PROTOCOL FOR ADHOC NETWORKS

The fundamental requirements for a secure routing protocol for adhoc networks are listed as below:

Detection of Malicious Nodes

• A routing protocol
→should be able to detect the presence of any malicious-node in the network &
→should avoid the participation of such nodes in the routing process

Guarantee of Correct Route Discovery

• If a route between the source & destination node exist, the routing protocol
→should be able to find the route &
→should also ensure the correctness of the selected route

Confidentiality of Network Topology

• Once the network topology is known, the attacker may try to study the traffic pattern in the network.
• If some of the nodes are found to be more active compared to others, the attacker may try to mount attacks.
• This may ultimately affect the ongoing routing process. Hence, confidentiality of network topology is important.

Stability against Attacks

• The routing protocols must be self-stable in the sense that it must be able to revert to its normal operating state within a finite amount of time after passive or an active attack.

SECURITY AWARE ADHOC ROUTING PROTOCOL (SAR)

• This uses security as one of the key metrics in path finding.
• In adhoc networks, communication between end-nodes through possibly multiple intermediate-nodes is based on the fact that the two end-nodes trust the intermediate-nodes.
• This defines level of trust
→as a metric for routing &
→as one of the attributes for security to be taken into consideration while routing

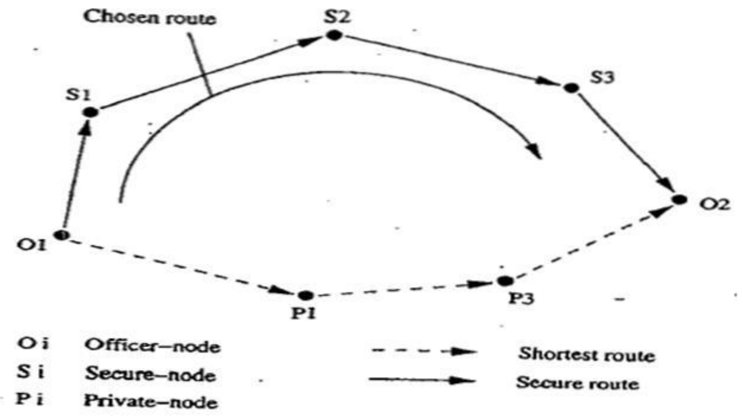


Figure 9.14. Illustration of the level of trust metric.

• Two paths exist between the two officers O1 and O2 who want to communicate with each other (Figure: 9.14),
• One of these paths is a shorter path which runs through private-nodes whose trust levels are very low
• Hence, the protocol chooses a longer but secure path which passes through other secure-nodes
• Nodes of equal levels of trust distribute a common key among themselves and with those nodes having higher levels of trust
• This could be incorporated into both on-demand and table-driven routing protocols
• This allows the application to choose the level of security it requires.
But the protocol requires different keys for different levels of security
This tends to increase number of keys required when the number of security levels used increase

SECURE EFFICIENT ADHOC DISTANCE VECTOR ROUTING PROTOCOL (SEAD)

- This based on the destination-sequenced distance vector (DSDV) routing protocol
- This is mainly designed to overcome security attacks such as DoS and resource consumption attacks
- The uses a one-way hash function and does not involve any asymmetric cryptographic operation

DISTANCE VECTOR ROUTING

- Distance vector routing protocols belong to the category of table-driven routing protocols
- Each node maintains a routing table containing the list of all known routes to various destination nodes in the network
- The metric used for routing is the distance measured in terms of hop-count
- The routing table is updated periodically by exchanging routing information
- An alternative approach to this is triggered updates, in which each node broadcasts routing updates only if its routing table gets altered.

ONE-WAY HASH FUNCTION

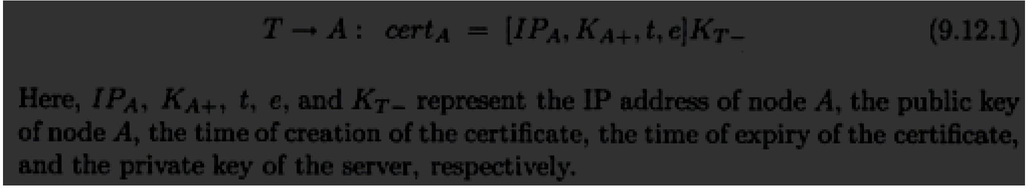
- SEAD uses authentication to differentiate between updates that are received from non-malicious-nodes and malicious-nodes.
- This minimizes resource consumption attacks caused by malicious-nodes.
- SEAD uses a one-way hash function for authenticating the updates.
- A one-way hash function (H) generates a one-way hash chain (h_1, h_2 ,......).
- The function H maps an input bit-string of any length to a fixed length bit-string.
- To create a one-way hash chain, a node generated a random number with initial value x € (0,1)p, where p is the length in bits of the output bit-string.
- h0 is the first number in the has chain is initialized to x.
- The remaining values are computed using a general formula $h_i = H(h_{i-1})$ for $0 \leq i \leq n$, for some *n*.
- SEAD avoids routing loops unless the loop contains more than one attacker.
- The protocol is robust against multiple coordinated attacks.
- SEAD protocol would not be able to overcome attacks where the attacker uses the same metric and sequence number which were used by the recent update message, and sends a new routing update.

AUTHENTICATED ROUTING FOR AD HOC NETWORKS (ARAN)

- It is a secure routing protocol which successfully defeats all identified attacks in the network layer.
- It takes care of authentication, message integrity and non-repudiation.
- During the route discovery process of ARAN, the source node broadcasts RouteRequest packets.
- Destination packets respond by uni-casting back a reply packet on the selected path.
- The ARAN protocol uses a preliminary cryptographic certification process, followed by an end-to-end route authentication process, which ensures secure route establishment.

ISSUE OF CERTIFICATES

- There exists an authenticated trusted-server whose public-key is known to all legal nodes in the network.
- The ARAN protocol assumes that keys are generated a priori by the server and distributed to all nodes in the network.
- On joining the network, each node receives a certificate from the trusted-server.
- The certificate received by a node A from the trusted-server T looks like the following:

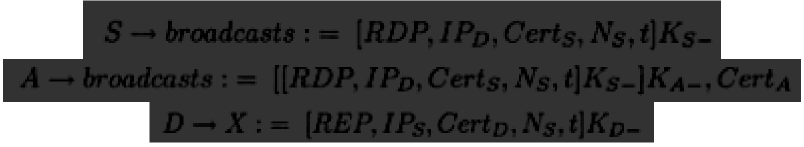


$$T \rightarrow A: \ cert_A = [IP_A, K_{A+}, t, e]K_{T-} \qquad (9.12.1)$$

Here, IP_A, K_{A+}, t, e, and K_{T-} represent the IP address of node A, the public key of node A, the time of creation of the certificate, the time of expiry of the certificate, and the private key of the server, respectively.

END-TO-END ROUTE AUTHENTICATION

- The main goal is to ensure that the correct intended-destination is reached by the packets sent from the source- node.
- The source-node *S* broadcasts a *RouteRequest/RouteDiscovery* packet destined to destination node *D.*

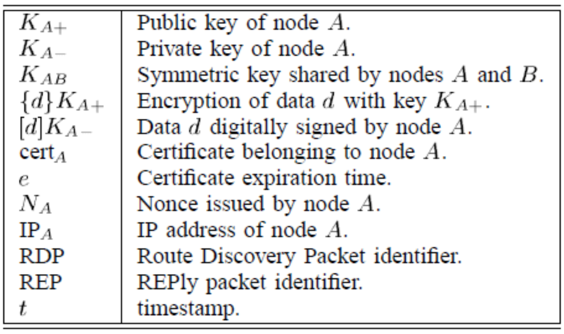


$$S \rightarrow broadcasts := [RDP, IP_D, Cert_S, N_S, t]K_{S-}$$

$$A \rightarrow broadcasts := [[RDP, IP_D, Cert_S, N_S, t]K_{S-}]K_{A-}, Cert_A$$

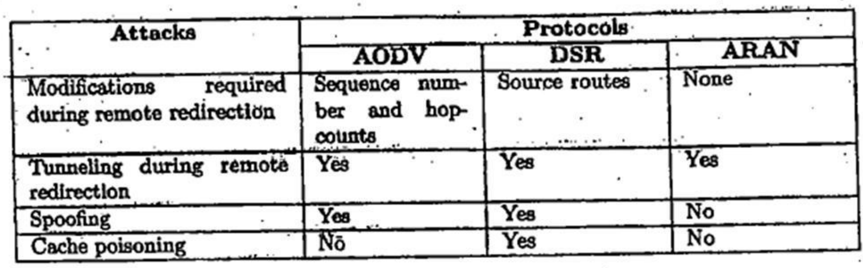
$$D \rightarrow X := [REP, IP_S, Cert_D, N_S, t]K_{D-}$$

Where,



K_{A+}	Public key of node A.
K_{A-}	Private key of node A.
K_{AB}	Symmetric key shared by nodes A and B.
$\{d\}K_{A+}$	Encryption of data d with key K_{A+}.
$[d]K_{A-}$	Data d digitally signed by node A.
cert$_A$	Certificate belonging to node A.
e	Certificate expiration time.
N_A	Nonce issued by node A.
IP$_A$	IP address of node A.
RDP	Route Discovery Packet identifier.
REP	REPly packet identifier.
t	timestamp.

Table 9.3. Comparison of vulnerabilities of ARAN with DSR and AODV protocols



Attacks	Protocols		
	AODV	DSR	ARAN
Modifications required during remote redirection	Sequence number and hop-counts	Source routes	None
Tunneling during remote redirection	Yes	Yes	Yes
Spoofing	Yes	Yes	No
Cache poisoning	No	Yes	No

SECURITY AWARE AODV PROTOCOL

- AODV is an on-demand routing protocol where the route discovery process is initiated by sending RouteRequest packets only when data packets arrive at a node for transmission.
- A malicious-node could advertise that it has the shortest path to the destination, thereby redirecting all the packets through itself. This is known as *blackhole attack* (Figure: 9.15),

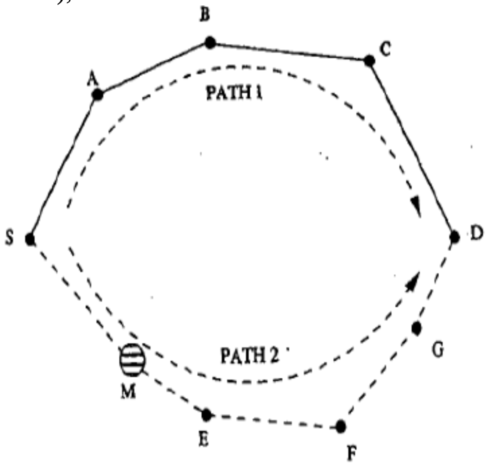


Figure 9.15. Illustration of blackhole problem.

- Let node M be the malicious-node that enters the network.
- It advertises that it has the shortest path to the destination-node D when it receives the RouteRequest packet sent by node S.
- The attacker may not be able to succeed if node A, which also receives the RouteRequest packet from node S, replies earlier than node M.
- Advantage: malicious-node does not have to search its routing table for a route to the destination.
- Hence the malicious-node would be able to reply faster than node A.

SOLUTIONS FOR THE BLACK HOLE PROBLEM

- One of the solutions for the blackhole problem is to restrict the intermediate-nodes from originating RouteReply packets (Figure: 9.16),
- Only the destination-node would be permitted to initiate RouteReply packets.
- Security is not completely assured.
- The delay involved in the route discovery process increases as the size of the network increases.

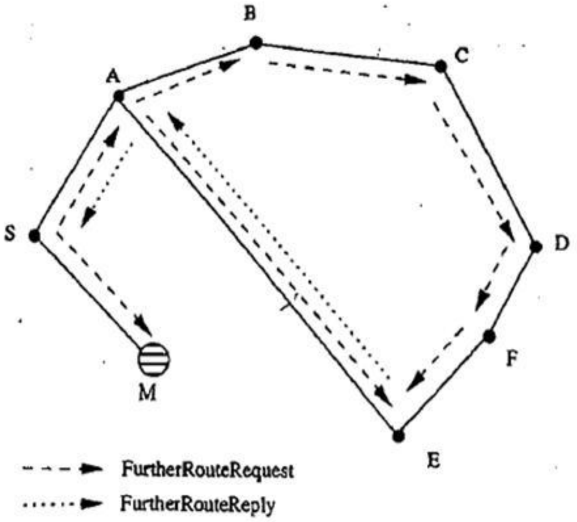


Figure 9.16. Propagation of *FurtherRouteRequest* and *FurtherRouteReply*.

- The source-node S sends FurtherRouteRequest packets to this neighbour-node E.
- Node E responds by sending a FurtherRouteReply packet to source-node S.
- Since node M is a malicious-node which is not present in the routing list of node E, the FurtherRouteReply packet sent by node E will not contain a route to the malicious-node M.
- This protocol completely eliminates the blackhole attack caused by a single attacker.
- Disadvantage: control overhead of the routing protocol increases considerably.
- If the malicious-nodes work in a group, this protocol fails miserably.